

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</small>				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 17 Jul.00		3. REPORT TYPE AND DATES COVERED DISSERTATION
4. TITLE AND SUBTITLE A QUANTITATIVE ANALYSIS OF THE EFFECT OF MARKET DESIGN AND POLICY UNCERTAINTY ON INVESTMENT IN ELECTRICITY GENERATION: A REINFORCEMENT LEARNING APPROACH			5. FUNDING NUMBERS	
6. AUTHOR(S) CAPT GROBMAN JEFFREY H				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) COLORADO SCHOOL OF MINES			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) THE DEPARTMENT OF THE AIR FORCE AFIT/CIA, BLDG 125 2950 P STREET WPAFB OH 45433			10. SPONSORING/MONITORING AGENCY REPORT NUMBER  FY00-251	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT Unlimited distribution In Accordance With AFI 35-205/AFIT Sup 1			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words)				
<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;"> <b>DISTRIBUTION STATEMENT A</b>            Approved for Public Release            Distribution Unlimited         </div> <div style="font-size: 2em; font-weight: bold;">20000727 172</div> </div>				
14. SUBJECT TERMS			15. NUMBER OF PAGES 168	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	

# A QUANTITATIVE ANALYSIS OF THE EFFECT OF MARKET DESIGN AND POLICY UNCERTAINTY ON INVESTMENT IN ELECTRICITY GENERATION: A REINFORCEMENT LEARNING APPROACH

By

Capt Jeffrey H. Grobman  
([njgrobman@earthlink.net](mailto:njgrobman@earthlink.net))

Evidence exists that electric market design and policy uncertainty significantly impact long-run electric generation investment. This research, which is organized in three separate essays, quantifies this relationship and in doing so provides policy makers with insights into the long-run implications of several proposed policies. It utilizes an innovative modeling technique, which has not previously been applied to this problem domain, to address the problem of modeling sequential investment under uncertainty.

The first essay introduces a general modeling framework that utilizes reinforcement learning (RL)—a recently developed technique for solving stochastic control problems—to model optimal long-run generation investment from both social welfare maximizing and monopolistic perspectives. This essay demonstrates that this technique can produce more realistic models of investment under uncertainty than other stochastic control methods because explicit definition of state transition probabilities is not required. Additionally, results show that models of generation investment that do not consider demand uncertainty may significantly over-predict investment levels due to the large up-front investment costs and per-period fixed costs associated with generation resources.

The second essay utilizes the framework presented in the first essay to determine the effect of capacity subsidies and price caps on investment and prices. Results show that capacity subsidies act to increase overall investment while reducing spot market price volatility. However, this policy increases total electricity prices once capacity charges are considered. Additionally, results show that the effects of spot market price caps differ based upon the modeling perspective. For the social welfare maximizer, higher price caps always lead to higher levels of investment, while the effects of price caps on average price are indeterminate. In contrast, for the monopolist, price caps produce an indeterminate effect on overall investment and prices are always equal to the cap.

The third essay uses the RL-based framework to investigate the manner in which policy uncertainty, relating to the enactment or repeal of investment tax credits (ITCs) and production tax credits (PTCs), impacts investment in wind power. Results show that the expectation of a potential ITC enactment may decrease the level of wind power investment due to the increased option value of waiting for the ITC. Expectation of a potential ITC removal may increase the rate of investment in wind power as firms speed up their rate of investment to take advantage of the ITC before it is removed. In contrast, expectation of a PTC will lead to an increase in wind power investment, and expectation of a PTC removal will result in a decrease in wind power investment. These differing responses to uncertain tax policy result from the fundamental characteristics of the policies. Those policies that reward firms based on the year of a specific investment will produce near-term investment results that are opposite in direction to the intended result of the proposed change. Also, since substitution opportunities exist between wind and classical technology investments, the investment postponing and enhancing effects of ITC expectation are stronger than those found in previous research.

A QUANTITATIVE ANALYSIS OF THE EFFECT OF MARKET DESIGN  
AND POLICY UNCERTAINTY ON INVESTMENT IN ELECTRICITY  
GENERATION: A REINFORCEMENT LEARNING APPROACH

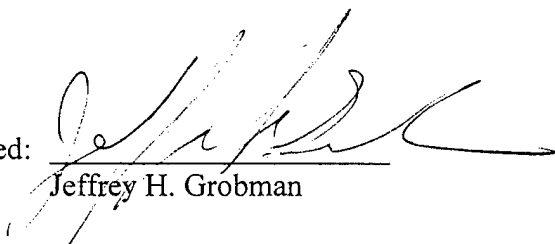
by

Jeffrey H. Grobman

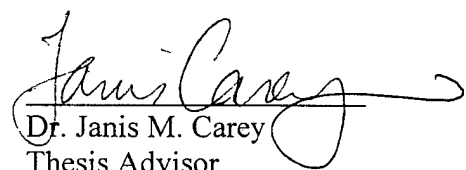
A thesis submitted to the Faculty and the Board of Trustees of the Colorado  
School of Mines in partial fulfillment of the requirements for the degree of Doctor of  
Philosophy (Mineral Economics).

Golden, Colorado

Date 5/26/00

Signed: 

Jeffrey H. Grobman

Approved: 

Dr. Janis M. Carey  
Thesis Advisor

Golden, Colorado

Date 5/26/00



Dr. Roderick G. Eggert  
Director, Division of  
Economics and Business

## ABSTRACT

Evidence exists that electric market design and policy uncertainty significantly impact long-run electric generation investment. This research, which is organized in three separate essays, quantifies this relationship and in doing so provides policy makers with insights into the long-run implications of several proposed policies. It utilizes an innovative modeling technique, which has not previously been applied to this problem domain, to address the problem of modeling sequential investment under uncertainty.

The first essay introduces a general modeling framework that utilizes reinforcement learning (RL)—a recently developed technique for solving stochastic control problems—to model optimal long-run generation investment from both social welfare maximizing and monopolistic perspectives. This essay demonstrates that this technique can produce more realistic models of investment under uncertainty than other stochastic control methods because explicit definition of state transition probabilities is not required. Additionally, results show that models of generation investment that do not consider demand uncertainty may significantly over-predict investment levels due to the large up-front investment costs and per-period fixed costs associated with generation resources.

The second essay utilizes the framework presented in the first essay to determine the effect of capacity subsidies and price caps on investment and prices. Results show

that capacity subsidies act to increase overall investment while reducing spot market price volatility. However, this policy increases total electricity prices once capacity charges are considered. Additionally, results show that the effects of spot market price caps differ based upon the modeling perspective. For the social welfare maximizer, higher price caps always lead to higher levels of investment, while the effects of price caps on average price are indeterminate. In contrast, for the monopolist, price caps produce an indeterminate effect on overall investment and prices are always equal to the cap.

The third essay uses the RL-based framework to investigate the manner in which policy uncertainty, relating to the enactment or repeal of investment tax credits (ITCs) and production tax credits (PTCs), impacts investment in wind power. Results show that the expectation of a potential ITC enactment may decrease the level of wind power investment due to the increased option value of waiting for the ITC. Expectation of a potential ITC removal may increase the rate of investment in wind power as firms speed up their rate of investment to take advantage of the ITC before it is removed. In contrast, expectation of a PTC will lead to an increase in wind power investment, and expectation of a PTC removal will result in a decrease in wind power investment. These differing responses to uncertain tax policy result from the fundamental characteristics of the policies. Those policies that reward firms based on the year of a specific investment will produce near-term investment results that are opposite in direction to the intended result of the proposed change. Also, since substitution opportunities exist between wind and

classical technology investments, the investment postponing and enhancing effects of ITC expectation are stronger than those found in previous research.

## TABLE OF CONTENTS

	Page
ABSTRACT.....	iii
LIST OF FIGURES .....	ix
LIST OF TABLES.....	xi
ACKNOWLEDGMENTS .....	xi
 Chapter	
1. INTRODUCTION .....	1
2. USING REINFORCEMENT LEARNING TO SOLVE FOR OPTIMAL ELECTRIC GENERATION INVESTMENT UNDER DEMAND UNCERTAINTY .....	7
2.1 Introduction .....	7
2.2 Modeling of Electricity Generation Investment.....	9
2.2.1 Theories of Investment .....	9
2.2.2 Quantitative Planning Models.....	13
2.3 Markov Decision Processes and Reinforcement Learning.....	15
2.3.1 Markov Decision Processes .....	15
2.3.2 Reinforcement Learning .....	18
2.4 General Modeling Framework .....	21
2.4.1 General Modeling Assumptions .....	23
2.4.2 RL Module .....	24
2.4.2.1 Indexed Sets .....	25

	Page
2.4.2.2 State Space .....	26
2.4.2.3 Action Space .....	27
2.4.2.4 Transition Probabilities.....	27
2.4.2.5 Reward Structure .....	28
2.4.3 Simulation Module.....	31
2.5 Demonstration of General Model.....	32
2.5.1 Current Market Description .....	32
2.5.2 Assumptions and Methods.....	33
2.5.3 Sample Results.....	38
2.6 Algorithmic Modifications .....	43
2.6.1 State Space Sweeping .....	43
2.6.2 Learning Rate Decay.....	45
2.6.3 Softmax Action Selection .....	48
2.6.4 Termination Criteria.....	50
2.6.5 Implementation .....	51
2.7 Conclusions .....	51
3. THE EFFECT OF MARKET DESIGN ON ELECTRICITY GENERATION INVESTMENT UNDER DEMAND UNCERTAINTY .....	55
3.1 Background and Motivation.....	55
3.2 The Restructured Environment .....	57
3.3 Models of Generation Investment Under Restructuring .....	60
3.4 The Effect of Capacity Subsidies on Generation Investment.....	61
3.4.1 Background on Capacity Subsidies and Reserve Requirements.....	61
3.4.1.1 California Market Design (Energy-Only Market) .....	68
3.4.1.2 PJM Market Design (Reserve Requirement) .....	69
3.4.2 Model of Capacity Subsidies .....	72
3.4.3 Capacity Subsidy Results.....	75

	Page
3.5 The Effect of Price Caps on Generation Investment .....	81
3.5.1 Why Price Caps May be Implemented.....	81
3.5.2 Modeling the Effects of Price Caps .....	83
3.5.3 The Effect of Price Caps: Results .....	85
3.6 Sensitivity of Peak Price to Demand Elasticity .....	91
3.7 Conclusions and Policy Implications .....	93
4. THE EFFECT OF UNCERTAIN TAX POLICY ON INVESTMENT IN WIND POWER.....	98
4.1 Introduction .....	98
4.2 Public Policy Pertaining to Wind Power .....	100
4.3 Literature on Investment Under Policy Uncertainty .....	107
4.4 Model.....	110
4.4.1 Assumptions.....	110
4.4.2 Model Structure .....	112
4.4.3 Data .....	116
4.4.4 Scenarios .....	117
4.5 Results .....	120
4.6 Conclusions and Policy Implications .....	128
5. CONCLUSIONS.....	133
REFERENCES CITED.....	138
APPENDIX A. C++ CODE FOR GENERAL FRAMEWORK .....	146

## LIST OF FIGURES

Figure	Page
1. Q-learning Algorithm.....	21
2. Overview of Modeling Framework .....	22
3. Optimal Policy ( $\sigma=0$ ).....	38
4. Optimal Policy ( $\sigma=300$ ).....	39
5. Optimal Investment in CC .....	40
6. Optimal Investment in CT .....	40
7. CT as a Percentage of Total Additional Capacity.....	41
8. Mean Additional Capacity by Market Structure and Level of Uncertainty .....	42
9. Modified Tabular Q-learning Algorithm .....	44
10. Q-values by Epoch.....	46
11. Learning Rate Decay Algorithm.....	47
12. Q-values vs. Policy Convergence .....	50
13. Short-Run Spot Market.....	59
14. No Market Clearing Price with Inelastic Demand.....	64
15. Welfare Loss from an Inefficient Reserve Requirement or Capacity Subsidy .....	66
16. Calculation of Reserve Requirement Using Price Responsive Demand Curve.....	74
17. The Effect of Capacity Subsidies.....	76

Figure	Page
18. Effect of Capacity Subsidies on Peak Price .....	77
19. Effect of Capacity Subsidies on Spot Price .....	78
20. Effect of Capacity Subsidies on Total Price .....	79
21. Effect of Capacity Subsidies on Investment Composition .....	80
22. Capacity Premium during Periods of Peak Demand.....	82
23. The Effects of a Price Cap .....	85
24. Effects of Price Cap on Monopolist.....	86
25. Effects of Price Cap on Social Welfare Maximizer.....	89
26. Effect of Price Caps on Price in Each Load Duration Curve Segment.....	90
27. Effects of Elasticity on Peak Price.....	92
28. Total Installed United States Wind Power Capacity.....	104
29. Varying ITC Level (Scenario 1) .....	120
30. Effects of Varying PTC Level (Scenario 2).....	121
31. Effects of Expected ITC Removal or Addition (Scenario 3A - 3D).....	122
32. Expectation of ITC Removal or Addition (Scenario 4A – 4D) .....	123

## LIST OF TABLES

Table	Page
1. Load Duration Curve Data.....	34
2. Technology and Cost Data.....	35
3. General Model Action Space .....	37
4. Run Times and Epochs of Learning by Scenario.....	51
5. Observed Capacity Values.....	73
6. Summary of Price Cap Effects.....	91
7. Action Space .....	114
8. Cost and Availability Data.....	116
9. Summary of Scenario Conditions.....	119
10. Mean Investment (MW) in Wind Power Across States without the ITC .....	125
11. Mean Wind Power Investment (MW) across States with the ITC.....	126
12. Mean Wind Power Investment (MW) across States without the PTC.....	127
13. Mean Wind Power Investment (MW) across States with the PTC.....	128

## ACKNOWLEDGMENTS

First, I would like to express my gratitude to my advisor, Dr. Janis Carey, for her guidance and mentoring throughout all phases of the dissertation process. The time and effort that Dr. Carey spent reading drafts and discussing this research is greatly appreciated. I am confident that the final product is an order of magnitude better than it would have been without her involvement.

Another professor that I must acknowledge is Dr. Manavendra Misra who introduced me to the field of reinforcement learning. Dr. Misra's work with me in the early stages of this research provided the foundation for the entire dissertation.

Additionally, I would like to thank the rest of my committee, Dr. R.E.D. Woolsey, Dr. Robert Frost, Dr. Wade Wartin, and Dr. William Navidi for their insightful comments and advice. I owe special thanks to Dr. Woolsey for exposing me to his unique outlook on the field of Operations Research. My interaction with him will impact me for the remainder of my career.

I also must recognize Mr. Morey Wolfson and Dr. Gary Schmitz from the Colorado Public Utilities Commission for giving me valuable insights into the workings of the electricity industry as well as providing me with the resources to complete this dissertation.

In addition, I owe thanks to my brother, Steve Grobman, who spent countless hours on the phone with me discussing the intricacies of the C++ programming language.

Finally, I would like to thank my loving wife, Norma, as well as my son, Tyler, for their support and love throughout this endeavor. On the many nights when I left Norma alone to care for our baby while I worked, I have no doubt that she had the more difficult job.

## Chapter 1

### INTRODUCTION

The United States electricity industry is in a period of structural change. Restructuring promises to bring competition to electricity generation. Additionally, environmental policies such as the Clean Air Act, the proposed renewable portfolio standard, and the Kyoto protocol may create new constraints for firms who wish to compete in this industry. Restructured electricity markets must be designed so that they promote competition and encourage technological innovation while maintaining the physical integrity of the system. This market design problem is nontrivial because of the technical complexities of electrical systems. Additionally, evidence from regions that have already restructured shows that firms will attempt to game any market rules that are established, and it is essential to update the market rules over time (Borenstein and Bushnell 1998; Green and Newberry 1992; Wolak and Patrick 1996; Wolak 1997). Also, firms may react to a policy, such as a proposed technology-specific subsidy, or to the prospect of a policy change, in a manner that is unintended by policy makers (Dixit and Pindyk 1994; Righter 1996).

Therefore, policy makers must carefully consider the short-run and long-run implications of any proposed policy prior to its implementation. Short-run policy concerns include mitigating market power and maintaining the physical security of the

electrical system. Long-run issues include ensuring adequate transmission and generation investment as well as providing for a socially optimal mix of generating resources. Most academic research has focused on short-run issues such as the mitigation of market power and less research has examined long-run concerns. This may be due to the fact that short-run policy issues are more pressing; however, the difficulty of modeling the long run may be another factor. The determination of the long-run effects of a proposed policy is difficult because both uncertainty and dynamics should be considered (Dixit and Pindyk 1994; McDonald and Siegel 1986; Pindyk 1991). Modeling electricity generation investment may be more complex than modeling investment in other industries because electricity demand varies throughout the year and firms can invest from a set of several generation technologies (Wang, Jaraiedi, and Torries 1996).

There is theoretical and empirical evidence that the regulated electricity industry did not motivate firms to invest in an efficient manner (Averch and Johnson 1962; Courville 1974; Gal-Or and Spiro 1992; Zajac 1970). Therefore significant opportunities for welfare gains may be possible through more efficient generation investment. Graves *et al.* (1998) estimate that potential cost savings from more efficient investment equal 10 to 15 percent. These levels exceed the potential cost savings from more efficient dispatch which are limited to 4 percent (Graves *et al.* 1998).

This research provides policy makers with a flexible framework for modeling generation investment that is capable of evaluating the long-run implications of proposed

policies. The framework is used to determine the effects of several specific policies on generation investment. This research is presented in three separate essays.

The first essay (Chapter 2) introduces a general modeling framework that utilizes reinforcement learning (RL)—a recently developed approach for solving stochastic control problems—to model optimal long-run generation investment from both social welfare maximizing and monopolistic perspectives. RL improves the ability of policy makers to analyze the long-run effects of proposed policies because it facilitates the development of realistic models that capture the effects of dynamics and uncertainty. Models utilizing classical stochastic control techniques, such as value iteration or policy iteration, have often lacked realism due to the curses of dimensionality and modeling. The curse of dimensionality refers to the exponential rise in computational time and memory required when computing a solution as the number of state and control variables increases (Rust 1996b). The curse of modeling refers to the inherent difficulty in explicitly defining all state transition probabilities (Bertsekas and Tsitsiklis 1996). Research in other problem domains has shown that RL can overcome these modeling difficulties and thereby produce rather realistic models of complex systems (Barto, Bradtke, and Singh 1991; Watkins 1989).

While RL has been applied to several other problem domains, it has not been used to model firm investment behavior. This research contributes to the literature by applying RL to the problem of modeling investment behavior. Also, this essay includes several novel modifications to the RL algorithm that facilitate this application. Results

show that RL can effectively model investment behavior and that techniques that do not explicitly consider uncertainty are prone to overestimate generation investment levels.

The second essay (Chapter 3) utilizes the RL-based framework presented in the first essay to examine the effects of capacity subsidies and price caps on generation investment and spot market prices. Capacity subsidies, or closely related reserve requirements, have been implemented in several restructured electricity markets to maintain system reliability and reduce price volatility (Singh and Jacobs 2000). Price caps have also been implemented in several markets to mitigate market power and to protect consumers from price spikes during peak demand periods (Wolak *et al.* 1999). These concerns of ensuring system reliability through markets, controlling price volatility, and controlling market power are relatively new issues for the electricity industry that were not present under the traditional system in which regulators dictated uniform reliability standards and regulated prices.

This analysis differentiates itself from other research because it approaches these issues quantitatively rather than qualitatively. Results show that capacity subsidies act to decrease price volatility by increasing the level of investment. However, capacity subsidies also increase the average total price of electricity, which includes the price of energy plus capacity payments. Price caps are also effective in reducing price volatility. Their impact on investment and average price varies based upon the market structure. An additional negative side effect of price caps is that they may force the system operator to shed loads to clear the market.

The third essay (Chapter 4) uses the RL framework to investigate the effect of policy uncertainty on investments in wind power. Specifically, the essay examines policy uncertainty relating to the enactment or repeal of investment tax credits (ITCs) and production tax credits (PTCs). Since the late 1970s, numerous policies such as ITCs and PTCs have been enacted at the state and federal level to promote investment in wind power as well as other renewable technologies (Cox, Blumstein, and Gilbert 1991). Investment in these technologies has been encouraged in order to offset investment in polluting fossil fuel-based technologies. An additional motivation for promoting renewable power is to develop a diverse fuel base for power production so that the economy is less vulnerable to the macroeconomic impacts of price shocks associated with one type of fuel. However, these state and federal policies toward wind power have changed regularly based upon the presidential administration, the composition of Congress, public attitudes toward renewable energy, and fossil fuel prices. Therefore, investors considering investment in wind power or other renewable energy technologies have faced considerable uncertainty over which policies will be in effect in the future.

This research contributes to the literature by analyzing the effects of policy uncertainty applied specifically to wind power investment. This class of problem is different from previous research on tax policy uncertainty which has considered only tax policies that apply to aggregate investment. In this research, tax policy uncertainty applies to only one technology from a larger group of substitutable technologies. Solution of this multi-technology model is facilitated by the RL modeling approach.

Results concur with those of Dixit and Pindyk (1994) and show that the expectation of an ITC can lead to a decrease in the rate of investment, whereas expectation of an ITC removal can result in an increase in the investment level. In contrast, the expectation of a PTC removal or addition will lead to a respective decrease or increase in investment.

## Chapter 2

# USING REINFORCEMENT LEARNING TO SOLVE FOR OPTIMAL ELECTRIC GENERATION INVESTMENT UNDER DEMAND UNCERTAINTY

### 2.1 Introduction

Reinforcement learning (RL) is a recently developed approach to solving infinite time horizon dynamic programming problems, often referred to as Markov decision processes (MDP) (Sutton and Barto 1998). Traditionally, this class of problems has been solved via well-established methods such as value iteration or policy iteration. However, these classical MDP solution techniques have difficulty addressing certain realistic problems due to the curses of dimensionality and modeling. The curse of dimensionality, applied to DPs, refers to the exponential rise in computational time and memory required when computing a solution as the number of state and control variables increases (Rust 1996b). The curse of modeling refers to the inherent difficulty in explicitly defining all state transition probabilities (Bertsekas and Tsitsiklis 1996).

RL, also known as neurodynamic programming, has shown promise to break the curses of modeling and dimensionality and thus facilitate modeling of larger and more complex problems (Barto, Bradtke, and Singh 1991; Watkins 1989). This is accomplished through an agent's "trial and error" interaction with its environment. RL algorithms have been applied successfully to several problem domains, including game

playing (Tesauro 1995), robotics and control (Connel and Mahadevan 1993), and dispatching problems (Crites and Barto 1996; Singh and Bertsekas 1997). The application of RL to economic problems has been scarce, but a few examples exist. Moody (1996) develops a RL-based model to develop optimal trading decisions and Van Roy (1998) uses RL to price high-dimension exotic derivatives.

This essay extends the reinforcement learning (RL) literature in two ways. First, RL is applied to a new application area, sequential investment behavior in uncertain environments. Additionally, several modifications to the basic tabular Q-learning RL algorithm are developed in order to facilitate the application of RL to the sequential investment problem domain.

Specifically, a general RL-based framework is introduced that determines investment level and technology choice decisions for electric generation investments from both social welfare maximizing and monopolistic perspectives. Next, this general model is demonstrated using the Rocky Mountain Power Area (RMPA) for differing levels of demand uncertainty.

The remainder of this essay is organized as follows: Section 2.2 provides background on the investment literature, and Section 2.3 provides detail on Markov decision processes and RL. Section 2.4 introduces a general RL-based framework for evaluating electric generation investment behavior, Section 2.5 applies this general framework to the RMPA, Section 2.6 discusses algorithmic developments for this implementation of RL, and Section 2.7 summarizes conclusions from this essay.

## 2.2 Modeling of Electricity Generation Investment

The complex technical realities of electrical power have motivated several planning models of electricity investment. This section first summarizes relevant general literature on modeling investment behavior in section 2.2.1 and then discusses specific models of investment that pertain to electricity generation in section 2.2.2.

### 2.2.1 Theories of Investment

The classical approach to modeling investment decisions, such as the decision to build a new electric power plant, is discounted cash flow analysis (DCF). This technique computes the present value of the expected cost of an investment and the present value of expected cash flows resulting from the investment. The differences in these values define the net present value of the investment and the investment is initiated if this value is positive (Stermole and Stermole 1996).

While NPV-based methods have served as the traditional means of modeling investment behavior, recent research has shown that they can produce severely biased results (Dixit and Pindyk 1994). This bias results from the failure of NPV analysis to consider the opportunity cost of investing rather than waiting for more information. The combination of irreversibility and uncertainty along with the ability to postpone an investment decision create this bias because if a firm decides to invest, it forgoes the opportunity to wait and learn more information about future realizations of uncertainty

(McDonald and Siegel 1986; Pindyk 1991). Therefore, generation investment decisions may be biased if NPV is used.

Another drawback with traditional NPV analysis is that it fails to consider managerial control once a project has been initiated (Smith and McCardle 1999). For instance, if a firm decides to invest in a coal-powered electrical plant, this plant may be shut down at some future date if unforeseen environmental regulations are enacted or if the price of coal rises to a point that makes the plant uneconomic.

In order to overcome these inadequacies, an options approach to investment has emerged which explicitly incorporates uncertainty and dynamics into the analysis. This approach is referred to as an options approach to investment because a firm can frame its investment decision as if it holds a financial call option. The firm may invest if it wishes, however, it is not obligated to do so (Dixit and Pindyk 1994).

Herbolet (1992) provides an empirical example of the importance of considering the option value of an investment by examining the decision of electric utilities to respond to Clean Air Act provisions regarding SO<sub>2</sub> emissions. To meet Clean Air Act standards, utilities can chose either to install scrubbers, switch to low-sulfur coal, or purchase tradable emissions credits. Results show that when an options approach is considered, purchasing credits may be preferable to the other alternatives, despite their lower NPV. This result arises from the added flexibility that purchasing credits provide because of their reversibility (Herbelot 1992).

One key difficulty in modeling firm behavior with the options approach is explicitly considering uncertainty and dynamics within the analysis. Available approaches that give explicit consideration to uncertainty are contingent claims analysis and decision analysis. Contingent claims analysis is implemented by replicating cash flows from an investment with a portfolio of tradable assets. Next, option valuation methods are used to value this portfolio of assets and develop the value of the investment (Dixit and Pindyk 1994, 94). A drawback to the contingent claims approach is that the technique assumes the stochastic component of assets that are being valued is perfectly correlated with that of a tradable asset (Dixit and Pindyk 1994, 121). This requirement may prohibit the modeling of certain complex problems, especially those involving entities that are not traded in markets. An advantage to the contingent claims approach is that a discount rate need not be defined exogenously, but rather is determined implicitly from market information (Dixit and Pindyk 1994, 121; Smith and McCardle 1999).

In contrast, decision analysis techniques such as stochastic programming and probabilistic dynamic programming require an exogenously defined discount rate but do not make any assumptions concerning tradable securities. Stochastic programming extends classical mathematical programming techniques to stochastic environments by enumerating possible future scenarios and then maximizing or minimizing the expected value of the objective function across scenarios (Birge and Louveaux 1997, 3). However, stochastic programming is severely limited due to computational difficulties when solving nonlinear, multi-stage, or discrete models (Birge and Louveaux 1997, 253).

In contrast, probabilistic dynamic programming (DP) is a much more general solution method, when compared with stochastic programming, and efficiently handles multi-stage or nonlinear problems (Rust 1996a). DP solves a larger problem by breaking it into a series of smaller problems via a series of backward recursions. Markov Decision Processes refer to infinite time-horizon dynamic programming problems and serve as the focus of this essay.

The research presented in this essay should not be confused with concurrent economic research by Erev and Roth in reinforcement learning which uses psychological learning theories to model the development of strategic behavior in repeated games (Erev and Roth 1998; Roth and Erev 1995). While similar, in that both areas of research involve reinforcement learning, two fundamental differences exist. Erev and Roth (1998) seek to understand how individual behavior evolves. They focus on human behaviors that are observed to be sub-optimal or irrational in experimental investigations. For example, experimental investigations in repeated games show that test subjects often act in a sub-optimal manner while improving toward optimal behavior with experience. Additionally, in some cases such as the ultimatum game, subjects exhibit behavior that diverges from optimality with experience (Roth *et al.* 1991). This use of RL contrasts with the application in this essay that uses RL to determine optimal or rational investment behavior. The second major difference between the approach used in this essay and the work of Erev and Roth deals with the structure of the RL model that is implemented. Erev and Roth (1998) update the probability that individuals will play a given game

theoretic strategy based upon reinforcement that they receive after playing that strategy. Therefore, if individuals receive a positive outcome, they are more likely to play a given strategy in the future. Similarly, if they receive a negative outcome, they are less likely to play a given strategy in the future. In contrast, the RL technique implemented in this essay maximizes expected discounted reward over an infinite time horizon. Therefore, in this essay's implementation of RL, a firm may choose to take an action with adverse immediate consequences provided that its long-run expected profit is maximized.

### 2.2.2 Quantitative Planning Models

The electric generation planning problem facing utilities is complex because they may invest in multiple technologies to meet widely varying demand. Other problem characteristics include the discrete nature of the control variables and uncertainty over future conditions at the time of the investment decision. Finally, reliability standards and pollution regulations further complicate the planning decision (Wang, Jaraiedi, and Torries 1996).

The complexity of this problem has motivated the development of many detailed cost-minimization models to aid regulated utilities with their planning. Additionally, these models have been used as positivistic tools for economists interested in predicting the way in which utilities will react to different types of environmental or regulatory restrictions (Wang, Jaraiedi, and Torries 1996).

Previous dynamic programming models of this problem include Booth (1972) and Levin, Tishler, & Zahavi (1983). Both of these efforts modeled the electric capacity expansion problem at the plant level. However, neither of these approaches considered uncertainty associated with future demands. Sherali and Soyster (1984) did account for the opportunity cost of waiting by using stochastic programming to address the capacity-planning problem. However, their model assumed that all variables were continuous—an assumption which clearly differs from the reality of lumpy capital in this industry.

One major difference between these models and the model presented in this essay is that this model assumes a profit or social welfare maximizing perspective rather than a cost minimizing perspective. The cost minimization framework may be appropriate for the regulated environment because franchised monopolies are obligated to serve all loads. Therefore, their objective is to minimize cost subject to this level of service constraint. This problem contrasts with the objective of a firm in a restructured environment whose goal is a maximization of profits.

One regulated scenario that would be modeled in a manner identical to the social welfare maximization perspective would be a dynamic cost of service case. In this type of regulation, a regulator forces a monopoly to make its investment decisions as well as short-run dispatch decisions in a manner that maximizes social welfare. It is important to note that this type of regulation assumes a great deal about the level of control that the regulator has over firm decisions compared with approaches such as rate-of-return regulation.

One additional difficulty in modeling the profit or social welfare maximization framework that is not present in the cost minimization framework is that the problem becomes nonlinear when demand is not perfectly inelastic. This occurs because price is a function of quantity and total revenues are calculated by multiplying price by quantity. This nonlinearity necessitates the use of dynamic programming in place of stochastic programming due to the problems associated with using stochastic programming to solve nonlinear models.

### 2.3 Markov Decision Processes and Reinforcement Learning

This section of the essay provides background on MDPs along with classical MDP solution techniques. Additionally, an explanation of the tabular Q-learning algorithm is provided.

#### 2.3.1 Markov Decision Processes

An MDP describes an agent which interacts with a system over a sequence of discrete time steps,  $t = 0, 1, 2, \dots, \infty$ . At each time step the agent is in a state  $s_t$  where  $s_t \in \mathcal{S}$ , the set of all possible system states. The agent then chooses an action  $a_t$  based upon its state where  $a_t \in \mathcal{A}(s_t)$ , the action space available to the agent from state  $s_t$ . Based solely on the state/action pair, the agent transitions to a new state  $s_{t+1}$ , or  $s'$ , based on the following transition probability:

$$P_{s,s'}^a = p\{s_{t+1} = s' | s_t = s, a_t = a\}. \quad (2.1)$$

Similarly, the agent receives a reward  $r_t$  whose expected value is based solely upon the state action pair:

$$R_{s,s'}^a = E\{r_t | s_t = s, a_t = a\}. \quad (2.2)$$

Equations (2.1) and (2.2) are critical criteria that must be met in order to have a Markovian system. If the path an agent takes to get to state  $s$  affects its transition probabilities or expected rewards associated with an action, the system cannot be modeled as an MDP (Sutton and Barto 1998).

Actions  $a_t$  are chosen based upon a policy  $\delta$  that maps states to actions  $\delta: S_t \rightarrow A_t$ . Specifically, each policy  $\delta$  determines the probability that action  $a_t$  will be chosen given that the agent is in state  $s_t$ :

$$\delta_s^a = p\{a_t = a | s_t = s\}. \quad (2.3)$$

Also, if the policy is not “mixed” (*i.e.*, only one action may be chosen for a given state) the policy may simply associate states with the indices of actions:

$$\delta_s = \arg(a_\delta). \quad (2.4)$$

The goal of an MDP system is to determine an optimal policy  $\delta^*$  which maximizes the expected discounted reward  $R_t$  for the system over an infinite time horizon. This reward is discounted by  $\gamma$  which is equal to  $1/(1+\text{discount rate})$ :

$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k} . \quad (2.5)$$

Many well-established DP approaches exist for solving MDPs including policy iteration, value iteration, generalized policy iteration, and linear programming (Ross 1982; Sutton and Barto 1998; Winston 1994). The majority of these techniques revolve around estimating value functions for each state. The value function for state  $s$  given policy  $\delta$ ,  $V_{\delta}(s)$ , is defined as the expected discounted reward given that the agent starts in state  $s$  and then follows the policy  $\delta$  from that point:

$$V_{\delta}(s) = E_{\delta}\{R_t \mid s_t = s\} = E_{\delta}\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s_t = s\right\} . \quad (2.6)$$

Value functions can be computed for any arbitrary policy  $\delta$  by solving the following set of recursive Bellman equations (Winston 1994):

$$V(s) = \sum_{s' \in S} P_{s,s'} (R_{s,s'}^a + \gamma V(s')) \quad \forall s \in S . \quad (2.7)$$

These Bellman equations are useful because expected rewards over an infinite time horizon can be expressed in two parts. These parts include the immediate reward and the expected discounted value across all successor states (Winston 1994, 1091).

Once optimal value functions are estimated, the optimal policy is simply the action that will transition into the successor state with the highest expected value. Thus, the literature often refers to the optimal policy as being “greedy” with respect to value. As long as immediate rewards are bounded, an optimal policy is guaranteed to exist (Winston 1994, 1091).

### 2.3.2 Reinforcement Learning

The RL algorithm utilized in this essay is the tabular Q-learning algorithm. Another RL approach is non-tabular Q-learning in which a function approximator is used to estimate Q-values based upon a set of features that define a state. The non-tabular approach is useful when the state space has too many dimensions to enumerate all possible combinations of features. Other variants of RL include SARSA, Actor-Critic methods, and Monte Carlo methods. Tabular Q-learning was selected from these methods because of its proven convergence properties and empirical evidence that it works well for many different types of problems (Jaakula, Singh, and Jordan 1994; Sutton and Barto 1998).

In the tabular Q-learning algorithm an agent interacts with its environment and, based upon the actions it selects, transitions from state-to-state. The analog to the value function in classical methods is the Q-value. State-action Q-values define the expected discounted reward if an agent starts in state  $s$  and then initially chooses action  $a$  but follows policy  $\delta$  from that point onward. The primary difference between Q-values and value functions is the addition of an initial action, which is independent of the policy under consideration. One can see this difference by contrasting equation (2.8) with equation (2.6) and noting that (2.8) includes the assumption that action  $a$  was chosen from state  $s$ :

$$Q_{\delta}(s, a) = E_{\delta}\{R_t \mid s_t = s, a_t = a\} = E_{\delta}\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s_t = s, a_t = a\right\}. \quad (2.8)$$

If state-action Q values are known, the optimal policy for a given state is the action with the largest associated Q-value.

$$\delta_s = \arg \max_a (Q(s,a)). \quad (2.9)$$

An  $\epsilon$ -greedy algorithm is one approach that is often used to select actions. This implies that the agent will select an action that is consistent with the agent's "current policy"  $(1-\epsilon)$  percent of the time. The current policy provides a state-to-action mapping based on equation (2.9). Using the current policy is defined as "exploitation." Occasionally, however, the agent will "explore" a new action that is chosen at random. The concept of combining both exploitation and exploration is critical for the convergence of most RL algorithms (Sutton and Barto 1998).

Another approach to action selection is a softmax algorithm that uses a Gibbs or Boltzmann distribution to select actions. With this approach, the probability of selecting action  $a$  is defined by:

$$p(a) = \frac{e^{Q_t(a)/\tau}}{\sum_{a \in A} e^{Q_t(a)/\tau}}, \quad (2.10)$$

where  $\tau$  is a "temperature parameter" and  $p(a)$  is the probability that an action from the set of actions  $A$  will be selected.

In this application, the distribution starts with a high temperature parameter  $\tau$  and then allows for cooling over time. This approach ensures that when the temperature is high, all actions have a near-equal probability of being selected. However, as the

temperature cools, those actions with higher  $Q$ -values have a higher chance of being selected. As was the case with the epsilon-greedy method, this approach to action selection balances exploration and exploitation.

The parameter in this distribution is labeled the temperature parameter because this distribution is used in the field of statistical mechanics to determine the probability that an atom will be in a given quantum energy state which relates to an atom's displacement from its ideal crystal position. As is the case with this application, when temperature is very high, the probability of an atom being in any given non-ideal location is equal. The temperature cooling in this application is somewhat analogous to the annealing process that involves the slow cooling of a metal. If a metal is cooled too quickly, its molecular structure will have imperfections as many atoms "freeze" in non-ideal locations. In contrast, if a metal cools slowly, the final product has fewer imperfections (Kittel 1996, 99).

After an action is chosen via an  $\epsilon$ -greedy or softmax approach, the agent receives a reward and transitions to  $s'$  where  $Q$ -values for state-action pair  $(s,a)$  are updated based on equation (2.11). One can observe that this algorithm only requires realizations of successor states  $s'$  and rewards  $r$  and thus it circumvents the "curse of modeling" because no explicit definition of transition probabilities and rewards is necessary. This characteristic has led some to classify this technique as a "model-free" method. The algorithm is summarized in Figure 1 (Sutton and Barto 1998).

```

Initialize Q(s,a)
Repeat for each episode
  Initialize s to s0
  Repeat for each step of episode
    Chose action (a) based on an action selection technique
    Implement action (a) and determine s' (the successor state) as well as the reward
    
$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[ r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \quad (2.11)$$

    s ← s'
  Until s is terminal
Until Q-values are sufficiently close to Q*

```

Figure 1. Q-learning Algorithm

In equation (2.11), the  $\alpha$  parameter is the learning rate and serves as the factor by which Q-values are adjusted following each iteration. Once optimal Q-values are found, the optimal policy  $\delta^*$  is determined based on equation (2.9).

#### 2.4 General Modeling Framework

The general modeling framework in this essay develops electric generation investment policies that maximize expected discounted monopoly profits or social welfare over an infinite time horizon in an uncertain environment. Additionally, this framework provides mean and variance information on investment and technology choice from any initial condition within the state space.

This framework can be applied to any region meeting the subsequently described assumptions of the model. Additionally, the framework can be modified by changing the reward structure, state space, or transition probabilities in order to evaluate policy issues relating to electrical generation investment. These analyses serve as the basis for

subsequent essays. The reinforcement learning approach facilitates this modeling flexibility because transition probabilities do not need to be defined for all states.

The framework incorporates basic assumptions on the nature of demand growth, technological parameters, as well as market structure to determine an optimal investment policy using RL. Next, similar assumptions are used in the simulation portion of the model along with initial state conditions to determine realizations of simulated investment outcomes. Both the MDP policy and simulation results provide insights into optimal investment behavior under varying conditions. Figure 2 summarizes this conceptual framework.

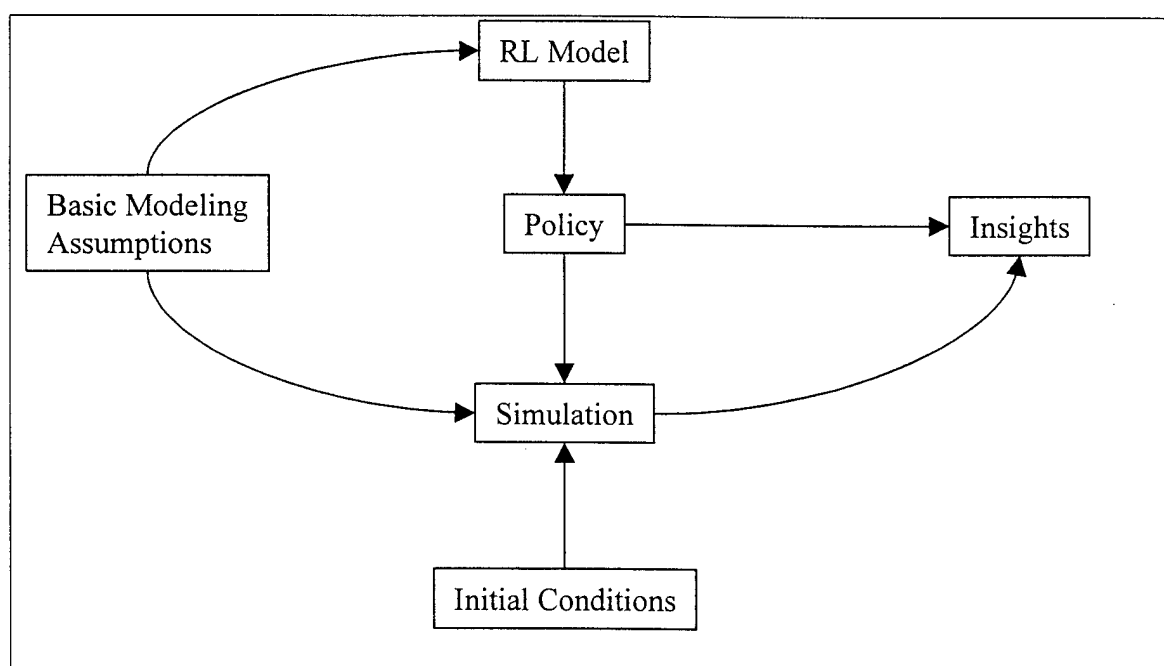


Figure 2. Overview of Modeling Framework

#### 2.4.1 General Model Assumptions

Discrete time. The model considers time discretely rather than continuously. Therefore, investment decisions may only be made at evenly spaced discrete points in time.

Investment Lead-times. The model assumes that investment lead-times of one period exist between the investment decision and the investment becoming “operational.” The agent has no knowledge of the stochastic component of demand growth between the investment decision and the subsequent period. The deterministic component of demand growth is known with certainty.

Irreversible and Bounded Investment. The model assumes that all investment is completely irreversible. Therefore, once a capacity investment is made, the agent is forced to pay fixed costs on this investment regardless of whether it is actually dispatched. Also, the model assumes that investment in each time period is bounded.

Transmission Constraints. The framework assumes that transmission within the region is unconstrained and produces no loss in load. Additionally, no access charges are accessed for the use of transmission.

Central Min-Cost Dispatch. The model assumes that generation units are dispatched based upon a min-cost dispatch from the lowest variable cost unit towards the highest variable cost unit until the desired total quantity of energy is dispatched.

Market Clearing. It is assumed that a regional market exists which determines a market-clearing price every load duration curve segment. This price is based solely upon

supply and demand bids. No provisions are made for bilateral contracts between generators and demanders. Additionally, no provisions exist for forward contracting or the use of other financial derivatives.

Load duration curve growth. The model assumes that the load duration curve shape remains constant from year-to-year. Additionally, it is assumed that the entire curve increases based upon a discrete state random walk with drift. Therefore, subsequent demand levels are independent of previous demand fluctuations. Finally, a discretized load duration curve is implemented as opposed to a continuous one.

Market Structure. Only monopolistic and social welfare maximizing perspectives are considered. No provisions exist for modeling cases of imperfect competition directly. However, the monopolistic and social welfare maximizing cases can be considered lower and upper bounds on investment resulting from imperfect competition.

Risk Neutrality. Neither maximization framework incorporates risk preferences other than risk neutrality.

No Externalities. The model assumes that no positive or negatives externalities exist relating to generation capacity or specific plant dispatch decisions.

#### 2.4.2 RL Module

The RL module determines an optimal policy mapping from the state space to the action space so that expected discounted rewards are maximized over an infinite time horizon. Since the rewards in this model represent yearly profits  $\Pi_t$  or social welfare  $SW_t$

for the monopolistic and social welfare maximizing perspectives respectively, the model maximizes:

$$\sum_{k=0}^{\infty} (\gamma^k \cdot \Pi_{t+k}), \quad (2.12)$$

or

$$\sum_{k=0}^{\infty} (\gamma^k \cdot SW_{t+k}), \quad (2.13)$$

where,  $\gamma$  is  $1/(1+\text{discount rate})$  for any  $t$ .

#### 2.4.2.1 Indexed Sets

In order to make the general model description clearer, the following indexed sets for time periods, technologies, and load duration curve segments are introduced. Time periods  $T = \{t | t=1, \dots, \infty\}$  signify the length of time between investment decisions. For example, if an investment were initiated in time period  $t=1$ , this investment would become operational in period  $t=2$ . The agent, either the social welfare maximizer or the monopolist, could then elect to initiate a new investment in  $t=2$  which would become operational in  $t=3$ . This period also identifies the “long run” because the capacity level may be adjusted over this time horizon. The classic definition of the long run in which all inputs are variable, is not met with this model because capacity can not be varied in an unconstrained manner over this time horizon. The set of available technologies  $H = \{i | i=1, \dots, M\}$  designates all of the generation technologies in which the agent may

invest. Similarly, the set of load duration curve segments  $J=\{j|j=1,\dots,N\}$  designates the set of loads that will be analyzed. This set is necessary when modeling electricity demand because demand curves are not static. Rather, market demand varies continuously over time.

#### 2.4.2.2 State Space

The agent's state  $S_t$  at time  $t$  is defined by a vector:

$$S_t = (D_t, K_{1,t}, K_{2,t}, \dots, K_{M,t}), \quad (2.14)$$

where  $D_t$  represents the value of the demand shift parameter at time  $t$  and  $K_{i,t}$  represents the capacity level for technology  $i$  in time period  $t$ . The demand shift parameter is a state variable that is multiplied by demand curves from all segments of the load duration curve.

The shift parameter  $D_t$  has an upper bound  $DMAX$ . Similarly, capacity levels for each technology  $i$  have an upper bound  $KMAX_i$ . The upper bound on demand  $DMAX$  is large enough so that it is outside of the relevant range of the model. Therefore, this bound will have an insignificant impact on investment decisions. This structure assumes that demand growth is independent of time over the model's relevant range. Upper bounds on capacity levels  $KMAX_i$  are set to accommodate a competitive dispatch if  $D_t$  were to equal  $DMAX$ . These upper bounds on  $D_t$  and  $K_{i,t}$  are necessary for application of the tabular Q-learning algorithm because Q-values for each state-action combination

must be stored in memory. If these values were unbounded, the state space would be of infinite size and the model would be intractable with the tabular Q-learning approach.

#### 2.4.2.3 Action Space

The action space is comprised of all combinations of vectors  $A_t$  representing investment in each of the  $M$  technologies during period  $t$ :

$$A_t = (I_{1,t}, I_{2,t}, \dots, I_{M,t}), \quad (2.15)$$

where  $I_{i,t}$  represents the quantity of investment in technology  $i$  during period  $t$ .

Values for  $I_{i,t}$  must be multiples of discrete values representing efficient plant sizes for each technology  $i$ .

#### 2.4.2.4 Transition Probabilities

The demand shift parameter evolves from year-to-year based upon a discrete state random walk with drift:

$$D_t = D_{t-1} + \theta + Z_t \quad \forall t \in T, \quad (2.16)$$

where  $\theta$  is a drift parameter representing growth over time and  $Z_t$  is a discretized normally distributed stochastic parameter.

Equations of motion for the capacity levels of each of the technologies can be represented by:

$$K_{i,t} = K_{i,t-1} + I_i \quad \forall i \in H, \forall t \in T. \quad (2.17)$$

Equations (2.16) and (2.17) are used to calculate state transitions unless  $D_t$  exceeds  $DMAX$  or  $K_{i,t}$  exceeds  $KMAX_i$  for technology  $i$ . If this occurs, the demand shift parameter or capacity level is set to its respective upper bound. A similar correction is applied if the model attempts to transition to a negative demand shift parameter value. Additionally, equation (2.16) is not utilized if demand is already equal to its upper bound. In this situation, the demand remains at the upper bound with a probability of 1. In this respect, the upper bound on demand acts as an absorbing boundary. These corrections ensure that all successor states are within the defined state space. It is necessary to define these boundary transitions despite their low likelihood of actual visitation, in order to implement the Q-learning algorithm that requires calculation of successor states from all states.

#### 2.4.2.5 Reward Structure of Monopolist and Social Welfare Maximizer

In general form, the monopolist's profits, not including fixed costs or investment costs, for segment  $j$  of the load duration curve in time period  $t$  is expressed by:

$$\pi_{j,t}(Q_{j,t}) = E_{j,t}(\cdot) - VC_{j,t}(\cdot) \quad \forall j \in J, \forall t \in T. \quad (2.18)$$

The social welfare maximizers reward, not including fixed costs or investment costs, for segment  $j$  and time period  $t$  is defined by:

$$sw_{j,t}(Q_{j,t}) = E_{j,t}(\cdot) - VC_{j,t}(\cdot) + CS_{j,t}(\cdot) \quad \forall j \in J, \forall t \in T, \quad (2.19)$$

where  $Q_{j,t}$  is the capacity that is dispatched in time period  $t$  and load duration curve segment  $j$ .  $E_{j,t}$ ,  $VC_{j,t}$ ,  $CS_{j,t}$  represent revenues from selling energy, variable costs, and consumer surplus during load duration curve segment  $j$  of time period  $t$  respectively.

$E_{j,t}$  represents revenues from the sale of energy and is defined as:

$$E_{j,t} = D_t^{-1} \cdot p_j(Q_{j,t}) \cdot Q_{j,t} \quad \forall j \in J, \forall t \in T, \quad (2.20)$$

where  $p_j(Q_{j,t})$  represents the inverse demand curve for load duration segment  $j$  and time period  $t$  and  $D_t^{-1}$  is the reciprocal of the demand shift parameter for period  $t$ .

Variable costs for each time period and load duration segment are represented by:

$$VC_{j,t} = \sum_{i \in H} q_{i,j,t} \cdot vc_i \quad \forall j \in J, \forall t \in T, \quad (2.21)$$

where  $q_{i,j,t}$  represents the quantity of energy produced by technology  $i$  in load duration segment  $j$  during time period  $t$  and  $vc_i$  represents the variable cost of producing 1 MWh of energy for technology  $i$ .

Consumer surplus for load duration segment  $j$  and time period  $t$  is defined by:

$$CS_{j,t} = \int_0^{Q_{j,t}} [D_t^{-1} \cdot p_j(Q_{j,t})] dQ - p_j(Q_{j,t}) \cdot Q_{j,t} \quad \forall j \in J, \forall t \in T. \quad (2.22)$$

Thus, total monopoly profits for time period  $t$  can be represented by:

$$\Pi_t = \sum_{j \in J} (s_j \cdot \pi_{j,t}(Q_{j,t})) - \sum_{i \in H} K_{i,t} \cdot fc_i - \sum_{i \in H} I_{i,t} \cdot ic_{i,t} \quad \forall t \in T, \quad (2.23)$$

and total social welfare for time period  $t$  can be expressed by:

$$SW_t = \sum_{j \in J} (s_j \cdot sw_{j,t}(Q_{j,t})) - \sum_{i \in H} K_{i,t} \cdot fc_i - \sum_{i \in H} I_{i,t} \cdot ic_{i,t} \quad \forall t \in T, \quad (2.24)$$

where  $s_j$  represents the percentage of the year for which load duration segment  $j$  is realized,  $ic_i$  represents investment costs for technology  $i$ , and  $fc_i$  represents the fixed costs associated with 1 MW of technology  $i$ . Equations (2.23) and (2.24) imply that fixed cost is only a function of capacity during time period  $t$  and not output.

Since the monopolist wishes to maximize its profits, the actual quantity of energy that is dispatched  $Q_{j,t}^*$  is defined by:

$$Q_{j,t}^* = \min \left[ \sum_{i \in H} K_{i,t}, \arg \max_Q \pi_{j,t}(Q) \right] \quad \forall j \in J, \forall t \in T. \quad (2.25)$$

This quantity is equal to the minimum of the monopolist's capacity and the quantity that maximizes its profits. In contrast, the social welfare maximizer dispatches  $Q_{j,t}^*$  MW of energy which is the minimum of capacity and the quantity that maximizes social welfare.

$$Q_{j,t}^* = \min \left[ \sum_{i \in H} K_{i,t}, \arg \max_Q sw_{j,t}(Q) \right] \quad \forall j \in J, \forall t \in T. \quad (2.26)$$

The actual levels of production  $q_{i,j,t}^*$ , for each technology  $i$  in load duration segment  $j$  during period  $t$  are selected based upon a min-cost dispatch of technologies from lowest to highest variable cost until  $Q_{j,t}^*$  is met. Therefore,  $q_{i,j,t}^*$  are determined based upon the following minimization:

$$\min_{q_{i,j,t}} \sum_{i \in H} q_{i,j,t} \cdot vc_i \quad (2.27)$$

subject to:

$$Q_{j,t}^* = \sum_{i \in H} q_{i,j,t}^* \quad \forall j \in J, \forall t \in T. \quad (2.28)$$

This minimization can be solved independently from the investment decision without sacrificing global optimality because this short-run dispatch problem is separable from the long-run investment decision.

#### 2.4.3 Simulation Module

The primary purpose of the simulation is to determine the mean and variance of generation capacity levels across time based upon an initial starting state. This type of information is critical because it is difficult to directly glean insights into how firms will actually invest from a multidimensional MDP policy. The simulation operates by making an agent's investment decision based upon the initial state of the system and the RL derived policy. This decision is then used, along with equations of motion (2.16) and (2.17) as well as the upper bounds on  $D_t$  and  $K_{i,t}$  to determine a subsequent state. This process is continued until time exceeds a predefined limit. Next, this process is repeated from the initial state until stable estimates for capacity means and variances as functions of time are developed. In the simulation module, all actions are selected based upon their Q-values so that for a given state, the action with the highest Q-value is always selected. This approach is used because it is assumed that the policy that is passed to the simulation from the RL module is optimal.

The framework can also be structured so a different stochastic parameter is used in the RL and simulation modules. This permits evaluation of the effects of misrepresenting uncertainty in policy formation. For instance, one could estimate the

losses that result from failing to consider uncertainty when formulating investment policy.

## 2.5 Demonstration of General Model

In order to demonstrate the previously described modeling framework, the framework is applied to the Rocky Mountain Power Area (RMPA). Optimal yearly investment policies are derived from both the monopolistic and social welfare maximizing perspectives for differing levels of demand uncertainty. Additionally, mean investment paths based on these optimal policies are also generated to illustrate the effect of market structure and uncertainty on investment behavior.

### 2.5.1 Current Market Description

The RMPA includes all of Colorado as well as eastern Wyoming. Electricity suppliers currently in this area include two investor-owned utilities (IOUs), twenty-six rural electric cooperatives, twenty-nine municipal utilities, and three joint action agencies (Sweetser 1998). The IOUs include West Plains Energy and Public Service Company of Colorado (PSCO) which is part of the holding company New Century Energies. PSCO possesses over 65 percent of the available generation capacity in the region (Sweetser 1998). Additionally, transmission capacity within the RMPA and between the RMPA and other surrounding regions is limited during peak hours.

Sweetser (1998) shows that the transmission restrictions of the RMPA combined with PSCO's large share of generation allow PSCO to exert market power, especially during peak load periods. Quick (2000) demonstrates that PSCO will have local monopoly power for up to 54 percent of the year as a result of transmission constraints.

### 2.5.2 Assumptions and Methods

First, it is assumed that all of the previously stated assumptions from Section 2.4.1 of the general model apply to the hypothetical region. Also, it is assumed that demand curves are iso-elastic and assume the form:

$$L_{j,t} = (D_t + D_j^0) \cdot p_{j,t}^\epsilon, \quad (2.29)$$

where  $L_{j,t}$  represents the quantity of energy demanded in load duration curve segment  $j$  of time period  $t$ .  $D_t$  is equal to the demand shift parameter and  $D_j^0$  is equal to the initial demand shift parameter level for each segment  $j$  of the load duration curve. Therefore, as  $D_t$  increases with time, the shape of the load duration curve remains unchanged while all demand curves shift outward. Values for  $D_j^0$  are set based upon the Borenstein, Bushnell, and Knittel (1999) "anchor point" method. For this technique, a reference price of \$30/MWh is chosen based upon the approximate electricity wholesale price in 1998 (Stone and Webster 1998). Next,  $D_j^0$  is varied until the quantity demanded matches the actual demand for this portion of the load duration curve. Table 1 summarizes the RMPA load duration curve data from 1998 that are used for these adjustments.

Table 1. Load Duration Curve Data

Index ( $j$ )	% of year ( $s_j$ )	Initial Load (MW) ( $D_j^0$ )
0	0.0001	4,000
1	0.0039	5,000
2	0.2029	6,000
3	0.2444	7,000
4	0.3188	8,000
5	0.1749	9,000
6	0.0501	10,000
7	0.0050	11,000

A price elasticity of demand  $\varepsilon$  of 0.1 is used. This estimate is within the range used by Borenstein, Bushnell, and Knittel (1999) who considered elasticities ranging from 0.1 to 0.4 in a recent market power study of California's restructured electricity market. A value from the lower range of reported elasticities was chosen because many consumers in the RMPA do not face real-time electricity prices and therefore have no demand-side response to price. This inelastic demand necessitates the use of a price cap for the monopolistic scenarios to prevent an infinite price markup. A cap of \$50/MWh is chosen arbitrarily. The implications of this choice are explored in Chapter 3. In addition, a trigger price of \$1000/MWh is set for calculation of consumer surplus to ensure that consumer surplus values are finite.

Combined cycle (CC) gas generation and combustion turbine (CT) gas generation are assumed to be the only available technologies for new investment. Capacity levels for these technologies along with various levels of the demand-shift parameter comprise the state space. These technologies are selected based upon the low cost of natural gas as

well as the environmental concerns associated with nuclear or coal powered plants. A similar assumption was made by a recent State-funded study investigating the effects of restructuring on the Colorado market (Stone and Webster, 1999). Also, over 99 percent of Colorado's capacity additions in 1998 consisted of either CC or CT units (DOE 1998). Table 2 contains cost data for these technologies. It is assumed that the quantity of each technology available for dispatch at any point in time is equal to 90 percent of the total installed capacity of that technology to account for scheduled and unscheduled maintenance. The model therefore assumes that plant availability does not vary with load.

Table 2. Technology Cost Data

	Combined Cycle	Combustion Turbine
Variable Cost ( <i>vc</i> )	17 \$/MWh	26 \$/MWh
Fixed Cost ( <i>fc</i> )	11,110 \$/MW	150 \$/MW
Investment Cost ( <i>ic</i> )	573,000 \$/MW	384,000 \$/MW

CC generation possesses significantly higher per-period fixed costs and up-front investment costs than CT generation while variable costs associated with CC generation are significantly lower than CT generation. This difference in cost originates from their designs. Combustion turbine generators operate similarly to a jet engine. They first utilize a compressor to compress incoming air. Next, this high-pressure air is mixed with gas in a combustion chamber. When the ignited gas passes out of the combustion chamber it turns a turbine which converts the thermal and kinetic energy into mechanical energy. This turbine is then used to generate electricity and the hot exhaust gases are

passed out. Combined cycle generation works similarly to the combustion turbine, however, the exhaust gasses are not wasted. This approach captures the exhaust gasses from the CT and uses them to power a steam turbine. The excess fixed and investment costs associated with CC generation result from the added steam recovery equipment as well as the additional steam turbine and generator. The fact that these exhaust gasses are recovered contributes to the higher efficiency and lower variable cost of the CC generator (GRI 2000).

The state space is designed so initial capacity is comprised solely of CC units. This assumption is made to account for the large quantity of low marginal cost coal plants currently in the RMPA.

State space capacity values in 150 MW increments range from 10,000 MW to 12,850 MW for CC, from 0 MW to 2,850 MW for CT generation and 0 MW to 2850 MW for the demand shift parameter yielding a total of 8000 states. A grid size of 150 MW is chosen because it is of sufficient fidelity to capture the dynamics of the problem while keeping run-times reasonable. Rust (1996a) discusses this trade-off between run-time and fidelity and suggests that even a relatively coarse grid is often sufficient to capture relevant economic phenomena. Another motivation for choosing 150 MW increments is that this value falls within the efficient plant size for both technologies that are under consideration (Fox-Penner 1997, 90). This state space is sufficiently large that upper bounds on the demand shift parameter have a negligible effect on the results.

The initial simulation state is determined so that the initial price is approximately equal to the 1998 average wholesale electricity price of \$30/MWh (Stone and Webster 1998). Simulation results are relatively invariant to this initial simulation condition after the first several simulated years. Also, the initial simulation state is set so that the initial value for the demand shift parameter is 900 MW greater than the lowest shift parameter value in the state space. This ensures that random downward fluctuations of the shift parameter will not be significantly affected by the lower “edge” of the state space.

The action space consists of the six actions listed in Table 3.

Table 3. General Model Action Space

Action Index	Investment in CC (MW)	Investment in CT (MW)
1	0	0
2	0	150
3	150	0
4	150	150
5	0	300
6	300	0

It is assumed that the drift parameter  $\theta$  is equal to 150 MW per year and the standard deviation of the stochastic parameter  $Z_t$  is equal to 150 MW based upon historical RMPA demand data from 1970-1998. However, in order to demonstrate the flexibility of the modeling framework, standard deviations for  $Z_t$  of 0, 150, and 300 are utilized in the RL module and a value of 150 MW is used in the simulation module for all cases. This allows for analysis of the effects of uncertainty on investment behavior. Finally, the discount parameter  $\gamma$  is set to 0.9 for all cases.

### 2.5.3 Sample Results

The optimal policies for the social welfare maximizer for standard deviations of  $Z_t$  of 0 and 300 appear in Figures 3 and 4 respectively. These figures show a fixed CC capacity of 10,000 MW for illustrative purposes. Similar graphs could be generated for all levels of CC capacity. In these figures, the capacity of CT generation is on the x-axis and demand shift parameter is on the y-axis. Total combined investment in both CC and CT generation is shown on the z-axis. The investment region appears in the upper-right-hand corner of the graph where capacity is low and demand is high.

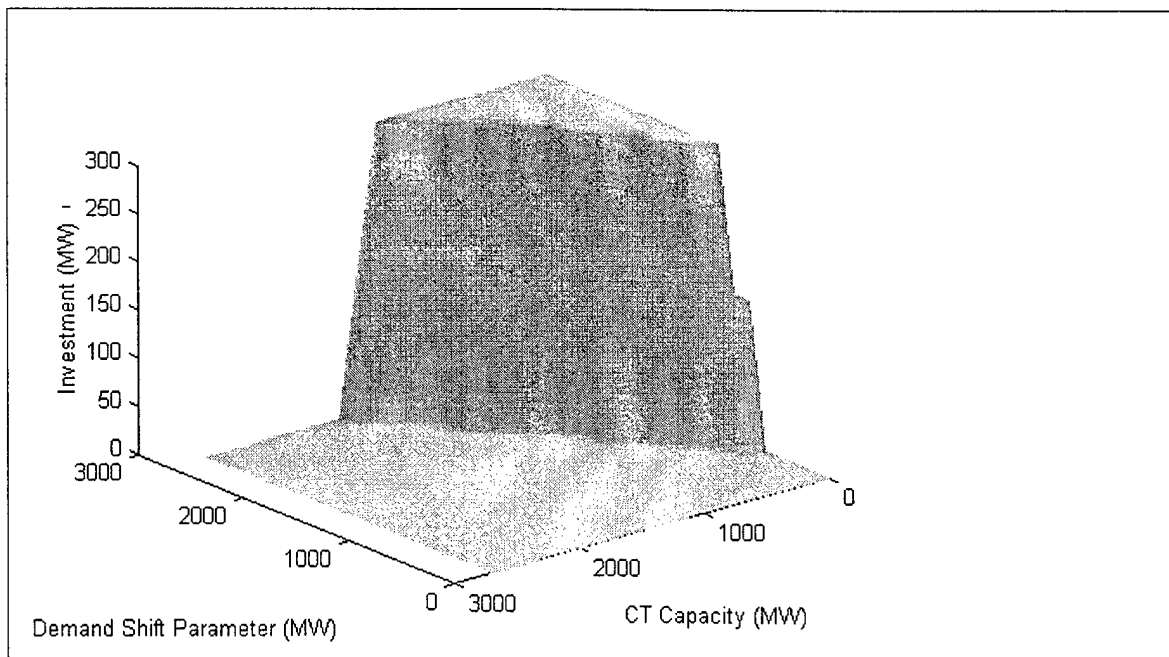


Figure 3. Optimal Policy ( $\sigma=0$ )

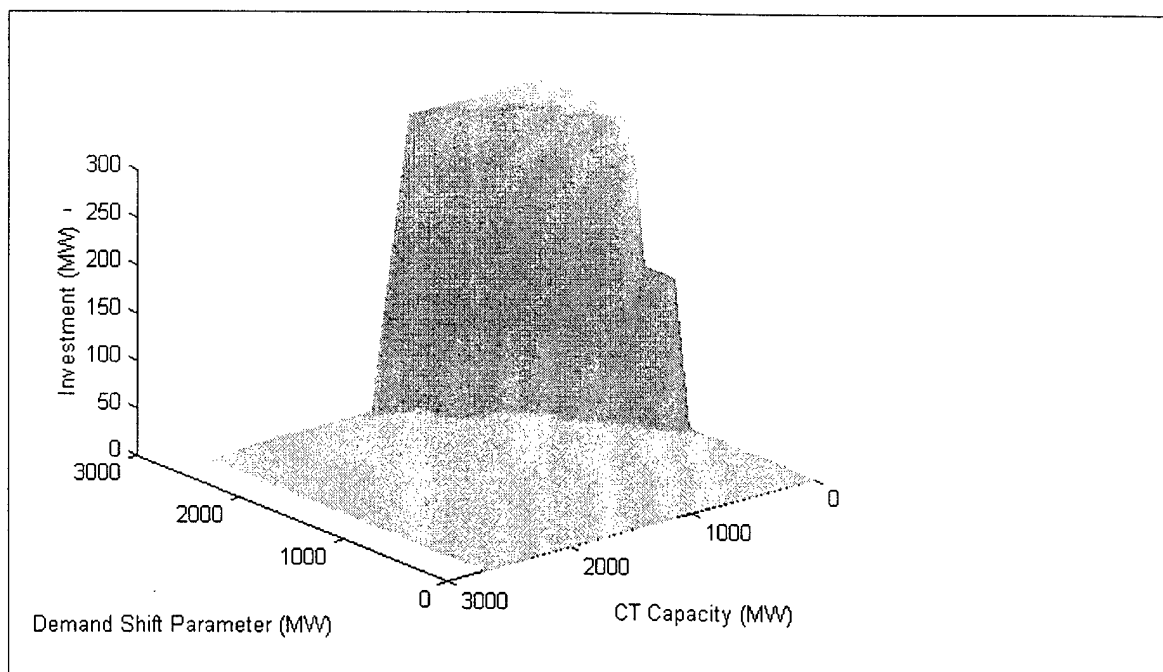


Figure 4. Optimal Policy ( $\sigma=300$ )

The investment region is slightly larger for the policy derived under certainty compared with the policy derived under uncertainty. Therefore, there are some states in which the agent facing certainty will invest and the agent facing uncertainty will not invest. This effect can be explained because of the option value of postponing the investment decision under uncertainty. Therefore, in the uncertain situation, the demand shift parameter must rise to a higher level prior to investment, compared with the certain situation.

Figures 5 and 6 decompose the graph showing total investment under certainty into investment by technology. CC and CT investments across the state space for the social welfare maximizer are shown in Figures 5 and 6 respectively.

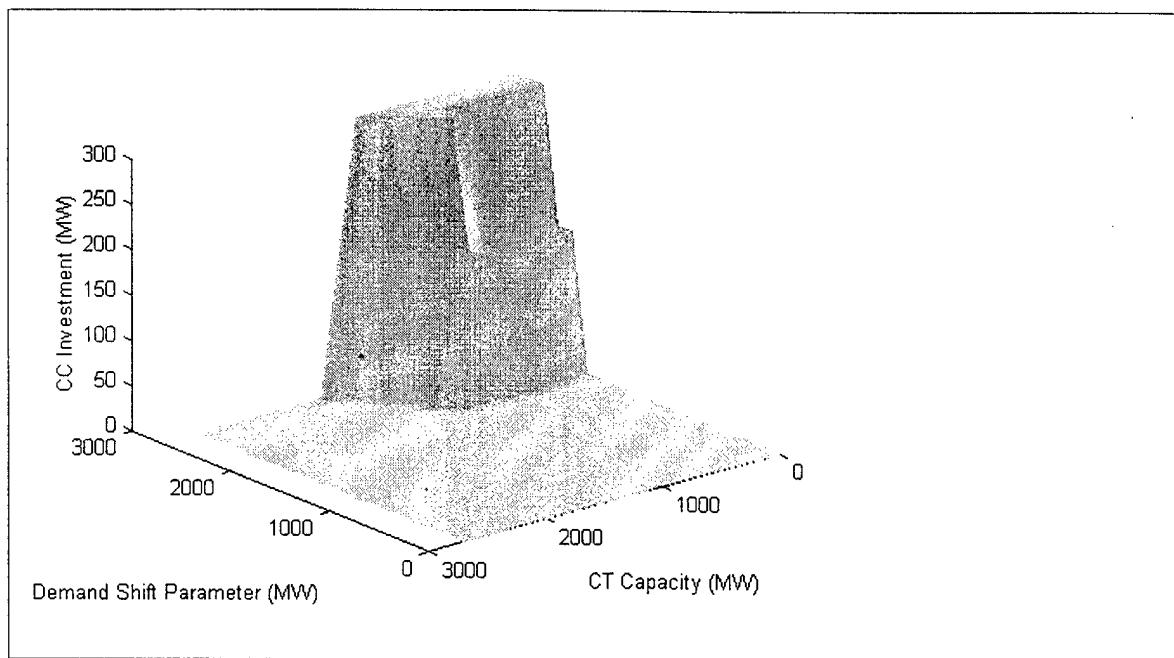


Figure 5. Optimal Investment in CC

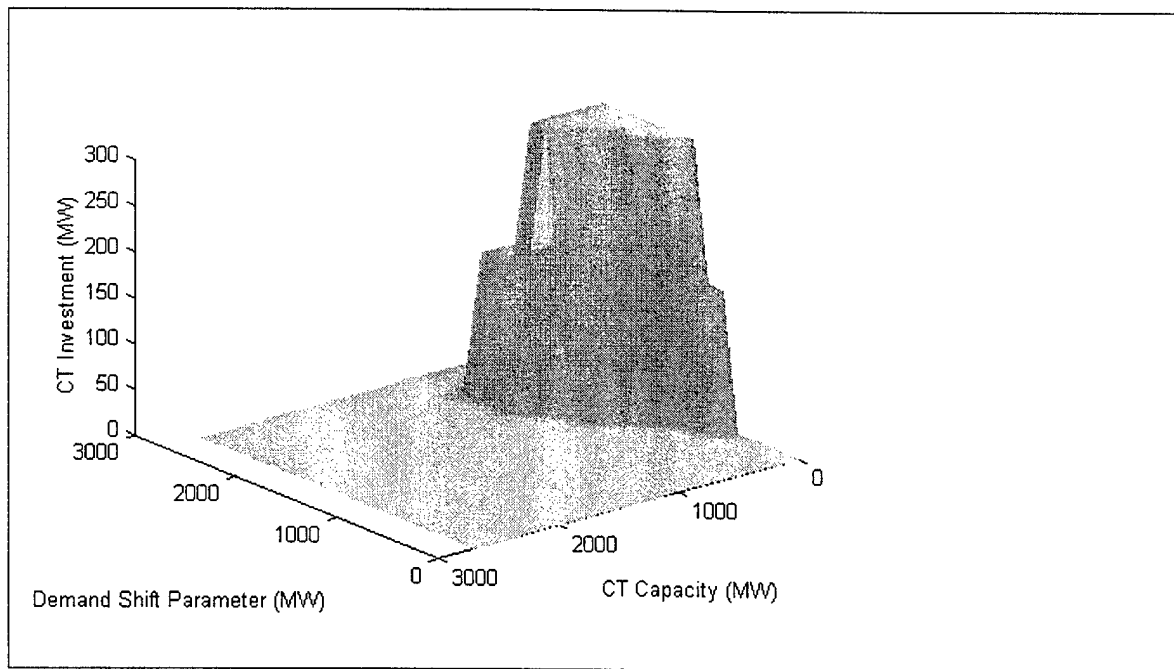


Figure 6. Optimal Investment in CT

The CC investment dominates when there is a large investment shortage, while, the CT investments are used to make up smaller shortfalls. This occurs because CC is used to meet all loads while CT investment only contributes to meeting peak loads. Similar results exist for the social welfare maximizer facing uncertain demand.

The implications of demand uncertainty on technology choice can be visualized in Figure 7 that plots the percentage of total additional capacity in year 15 that is comprised of CT units for each level of uncertainty.

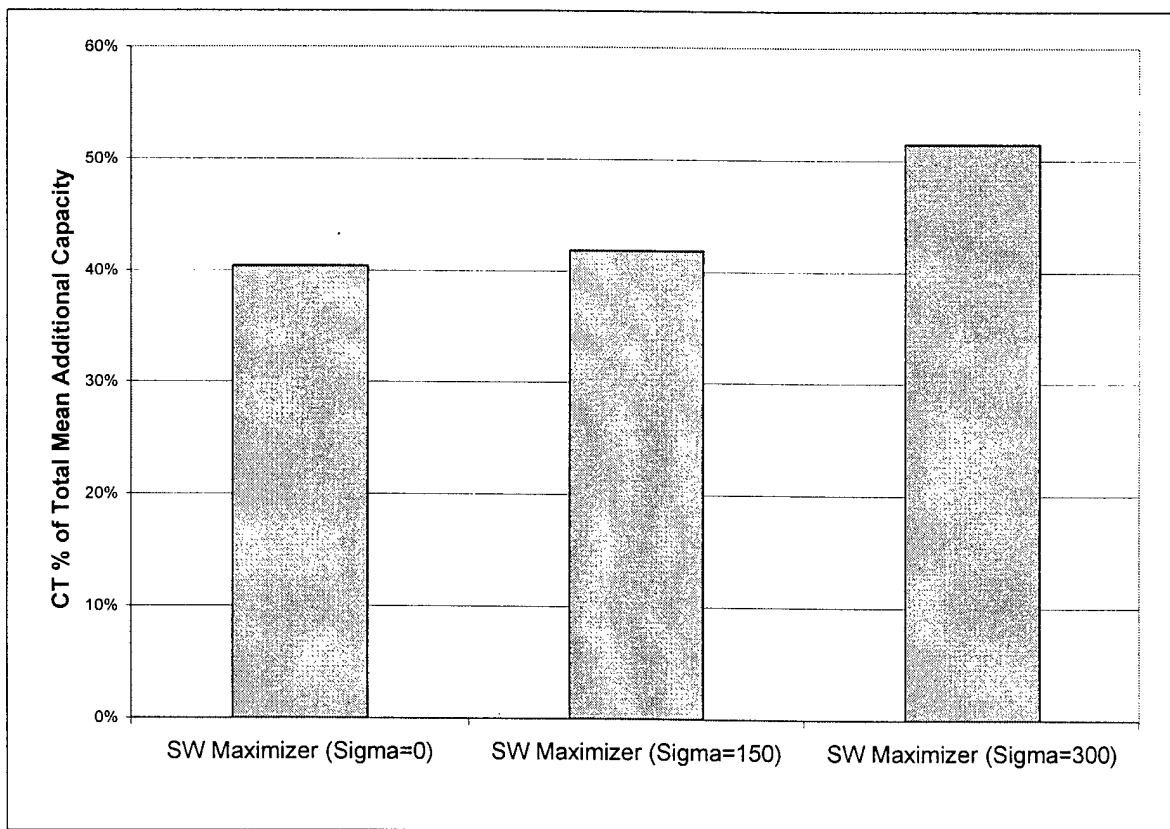


Figure 7. CT as a Percentage of Total Additional Capacity

Similar results exist for the other years. This graph illustrates that increased uncertainty causes the agent to prefer CT generation due to its lower fixed cost.

Figure 8 shows total mean additional capacity from the monopolistic and social welfare maximizing perspectives for varying levels of demand uncertainty. The social welfare maximizer invests at a higher average level compared with the monopolist as would be expected. Since, it is assumed that no externalities or distortionary taxes exist, the social welfare maximization scenario can be used to back out a perfectly competitive outcome (Dixit and Pindyk 1994, 283).

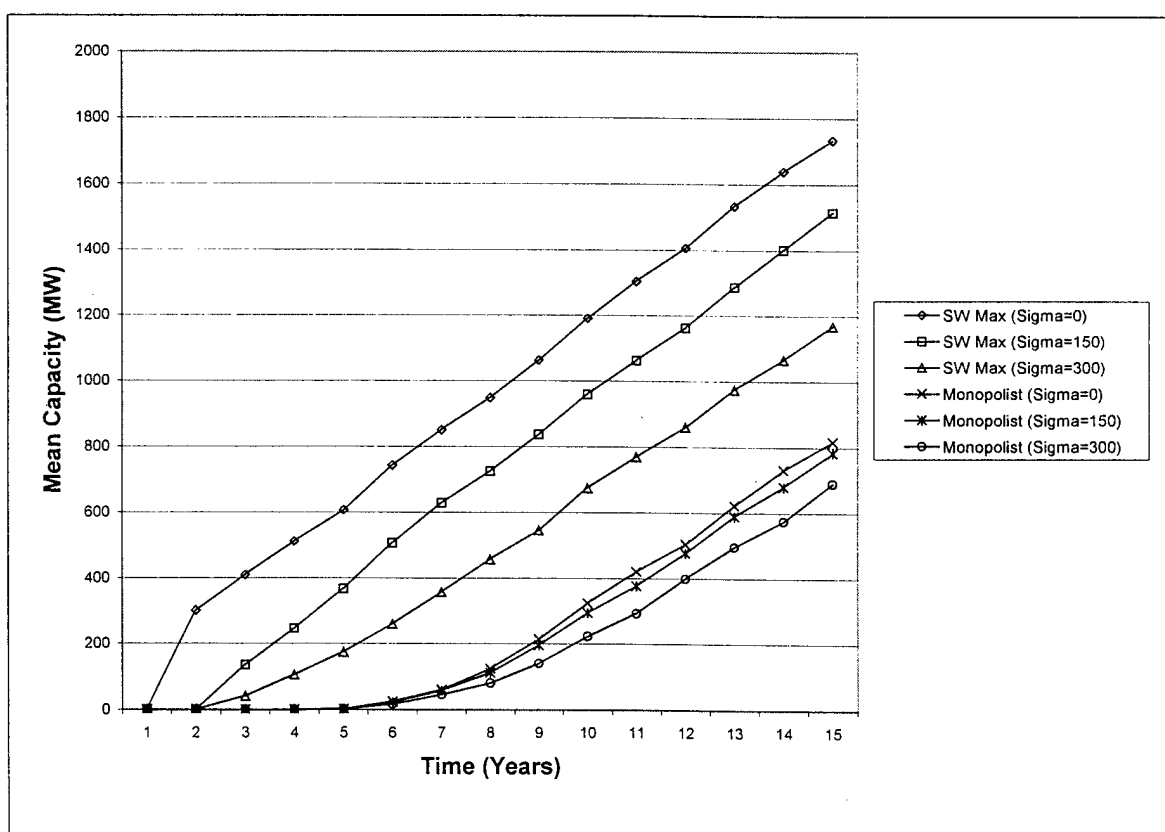


Figure 8. Mean Additional Capacity by Market Structure and Level of Uncertainty

Therefore, the monopoly and social welfare maximizing scenarios can be used to bound the level of investment resulting from a case of imperfect competition for each level of uncertainty. As expected, investment decisions that are formulated under higher levels of uncertainty result in reduced total capacity levels.

## 2.6 Algorithmic Modifications

Unlike many normativistic applications of RL which only require “good” solutions, it is essential to achieve near-optimal solutions with this model. This is necessary because this framework is designed for policy analysis in which one must compare results among differing policy alternatives. Therefore, solutions which are significantly suboptimal may misrepresent the effects of certain policies or in certain cases produce results which are opposite in sign to the actual underlying policy effect. In order to achieve optimal results in reasonable time periods, several modifications to the basic tabular Q-learning algorithm are made. This section of the essay summarizes these modifications and discusses state-space sweeping in 2.6.1, learning rate decay in 2.6.2, softmax action selection in 2.6.3, termination criteria in 2.6.4, and implementation in 2.6.5.

### 2.6.1 State Space Sweeping

Initially, when states were chosen for evaluation based upon the classical tabular Q-learning algorithm that is summarized in Figure 1, certain states were visited so

infrequently that the model could not learn an optimal policy in tractable run-times (less than 24 hours). This problem was overcome by forcing the Q-learning algorithm to evaluate each state by systematically sweeping the entire state space and executing one iteration of Q-learning on each state. Figure 9 summarizes the revised algorithm.

```

Initialize Q(s,a)
Repeat
  Evaluate sum of delta Q to consider learning rate decay
  For each s ∈ S
    Chose action (a) based on a softmax distribution
    Implement action (a) and determine s' (the successor state) as well as the reward
    
$$Q(s,a) \leftarrow Q(s,a) + \alpha [r + \gamma \max_{a'} Q(s',a') - Q(s,a)] \quad (2.30)$$

  End for
Until Q-values are sufficiently close to Q*

```

Figure 9. Modified Tabular Q-learning Algorithm

This algorithm is conceptually similar to one presented by Sutton and Barto (1998, 229) in which states are selected randomly from the state space after which one iteration of Q-learning is performed. However, random state selection did not work as well as systematic sweeps because the random approach did not visit certain states often enough to compute good Q-value estimates. Another key advantage of systematic sweeps is that the final policy for all states can be graphed to provide insights into optimal investment behavior. Graphs of optimal policies would not be meaningful with the classical implementation of Q-learning due to the poor accuracy of the policy at low-probability states.

It is important to note that this method of implementing tabular Q-learning does not take advantage of one of the key strengths of reinforcement learning, namely, the

ability to ration computational time to states based upon the probability that they may actually be visited. However, since run times were still reasonable, the computational inefficiency of the state space sweeping approach was not a serious obstacle.

### 2.6.2 Learning Rate Decay

One initial observation when working with the tabular Q-learning algorithm with fixed learning rates is that higher learning rates yield more rapid initial Q-value convergence compared with lower rates. However, larger learning rates tend to oscillate around their optimal value following convergence, which leads to sub-optimal estimation of the policy. This outcome contrasts with smaller learning rates that do not oscillate significantly around their optimal values but require significant time to converge. Figure 10 demonstrates this effect by comparing Q-values associated with 6 actions by epoch for learning rates equal to 0.5, 0.05, and 0.005. This figure is organized with epoch on the x-axis and Q-values on the y-axis. For illustrative purposes, the Q-values in this example are initialized to be close to the optimal Q-values, thus, allowing for rather rapid convergence.

While many other RL implementations use a “small” fixed learning rate, this approach is unacceptable for this application because learning rates that are small enough to provide sufficient accuracy, often result in intractable run-times (greater than 24 hours).

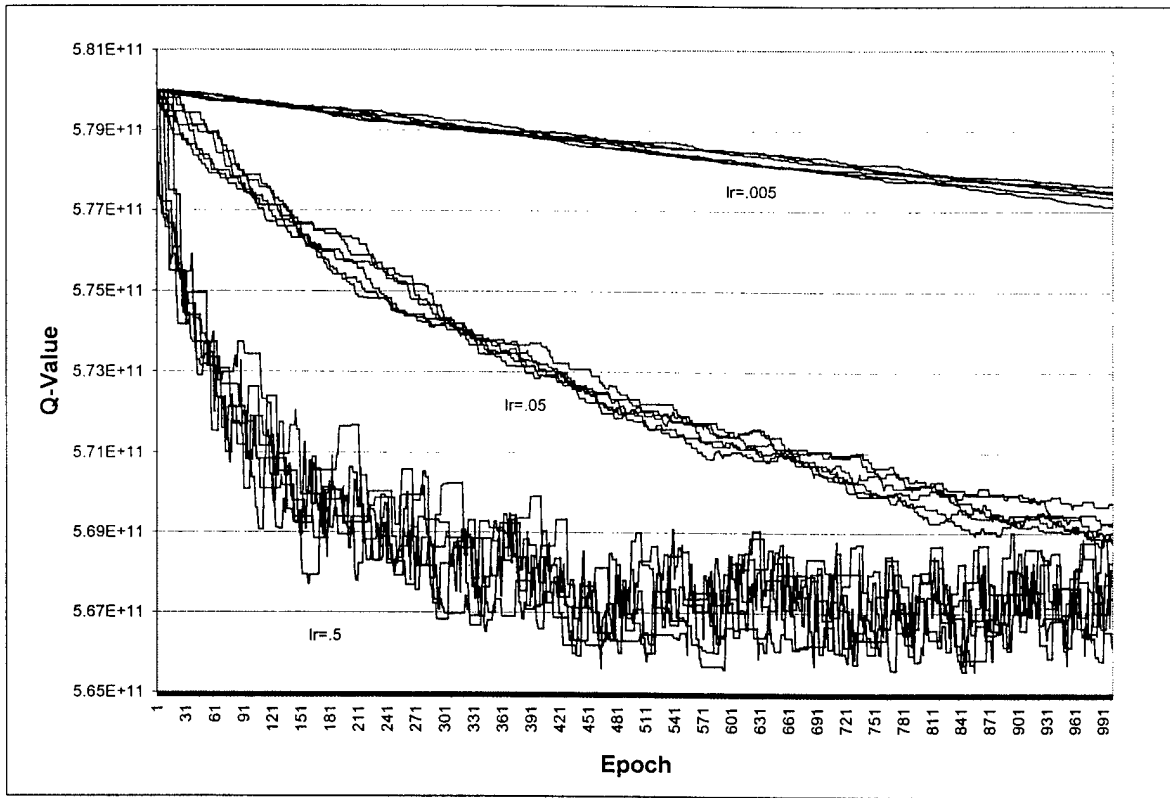


Figure 10. Q-values by Epoch for Varying Learning Rates

Because the literature does not provide much guidance concerning the design of learning rate decay algorithms, the following algorithm was developed. This algorithm, which is summarized in Figure 11, allows for learning rate decay so that initial convergence is rapid with a high learning rate. However, as learning progresses, the learning rate decreases to allow for a more accurate estimation of final Q-values. This in turn leads to a more accurate estimation of the optimal policy.

The algorithm keeps track of the sum of the absolute deviations in Q-values across the entire state space over every k-epoch period. When this value increases across

successive k-epoch periods, the learning rate is decreased and the process is repeated.

When the learning rate is decreased, it decreases based upon the following geometric series:

$$\alpha_n = \psi^n, \quad (2.31)$$

where,  $\alpha_n$  is the learning rate and  $n$  is the counter that is incremented every time the number of policy changes increases over a k-epoch period.

Intuitively, this algorithm is effective because initially a learning rate is effective in improving the policy through updating all of the Q-values in the state space. As learning progresses for a given learning rate, the absolute deviation in Q-values decreases as a given learning rate becomes less effective in improving the policy. Once the sum of the absolute deviations in Q-values increases, it is a sign that the current learning rate is not improving the policy and it is necessary to reduce the learning rate further.

```

Initialize  $\alpha$ 
Initialize  $\Delta_0$  to a large number
Initialize n
Repeat
    Run RL model for  $k$  epochs
     $\Delta_1 \leftarrow$  Sum of the absolute deviations in Q values across all states for  $k$  epochs
     $\Gamma \leftarrow (\Delta_0 - \Delta_1)$ 
    If  $\Gamma < 0$ 
         $n \leftarrow (n+1)$ 
         $\alpha \leftarrow (\psi)^n$ 
     $\Delta_0 \leftarrow \Delta_1$ 
Until termination criteria is met
  
```

Figure 11. Learning Rate Decay Algorithm

Based upon experimentation, a value for  $k$  of 100,000 was selected. Results show that smaller values decrease the learning rate too rapidly because they result in spurious increases in the number of policy changes. In contrast, larger sampling periods did not suffer from this drawback but did increase run times. Similarly, the geometric series with  $\psi$  equal to 0.1 was chosen based upon experimentation with different series. Experimentation also shows that  $\psi$  and  $k$  can be traded off against one another. This implies that a lower value for  $k$  necessitates a higher value for  $\psi$ . In no way is it suggested that this sampling period or this geometric series maximize the rate at which an optimal policy may be found; however, this algorithm does provide reasonable solutions for the problem under investigation.

This learning rate decay algorithm was not necessary when examining cases that did not involve uncertainty. For these cases, a fixed learning rate of 0.5 was used throughout the learning process.

### 2.6.3 Softmax Action Selection

Softmax action selection is implemented rather than the  $\epsilon$ -greedy approach. The drawback with  $\epsilon$ -greedy action selection is that it chooses actions other than the one with the highest Q-value with the equal probability  $\epsilon/(n-1)$ . For this application, there are usually several actions that are close to being the best action and some that are far from optimal. Therefore, an ideal action selection algorithm should focus on evaluating the

better actions while spending little computational time on those actions with lower Q-values. However, this preference for high Q-value actions should not be initiated until reasonable estimates for Q-values are attained. Since the softmax action selection algorithm incorporates this property, run times are significantly lower with this approach compared with the  $\varepsilon$ -greedy technique. There are situations in which the  $\varepsilon$ -greedy approach would be preferable to the softmax algorithm such as where only one action is a clear “winner” for each state and other actions are almost “equally bad.”

The drawback with softmax action selection is that if the temperature cools too rapidly, the algorithm may prematurely exclude certain actions that are in fact optimal. This problem is avoided by choosing a temperature cooling rate such that all actions across all states are chosen with a probability of at least 0.1 at the terminal epoch.

The cooling rate is determined based upon the exponential decay:

$$\tau_n = e^{-\rho \cdot n}, \quad (2.32)$$

where,  $\tau$  is the temperature at epoch  $n$  and  $\rho$  is a constant derived from:

$$\rho = \frac{-\ln(\tau_N)}{N}, \quad (2.33)$$

where,  $\tau_N$  is the desired temperature at the terminal epoch  $N$ . Various values for  $\tau_N$  were experimented with until one was found which ensured that all actions across all states were sampled with a probability of at least 0.1 in the terminal epoch.

#### 2.6.4 Termination Criteria

In order to determine when to terminate the Q-learning algorithm, the policy is monitored rather than the set of Q-values. This is done because the critical model output is an optimal policy rather than an optimal set of Q-values. Also, in certain applications there is evidence that an optimal policy is reached long before Q-values are near-optimal (Sutton and Barto 1998, 108). This principal is illustrated in Figure 12 which shows notional Q-values for two actions plotted by epoch. One can see that the optimal policy of action A is reached is reached long before Q-values approach their true optimal level.

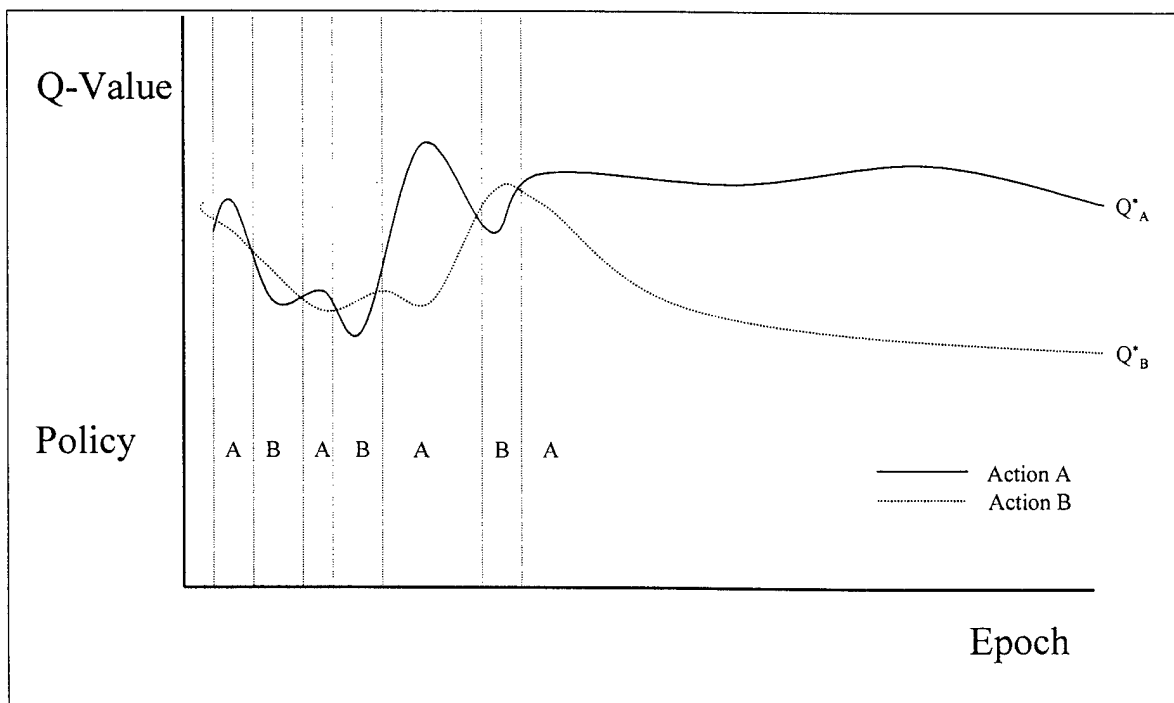


Figure 12. Q-values vs. Policy Convergence

Therefore, learning is terminated after the model completes 500,000 epochs without a change in policy. This heuristic was compared against much longer runs and in all cases the 500,000 epoch policy change test was more than sufficient to provide optimal results. Also, as was the case with learning rate decay, this termination criteria was unnecessary for cases that did not involve demand uncertainty. For these cases, learning was terminated once there were no policy changes after 100,000 epochs.

### 2.6.5 Implementation

The 8000 state RL model described in this essay was programmed in Microsoft Visual C++ © version 6.0. Run-times for this model varied widely depending upon the variance of the stochastic parameter with longer run-times associated with higher stochastic parameters. Table 4 summarizes the run-times required on a 400 Mz Pentium II for the cases considered. The C++ code for this model is contained in Appendix A.

Table 4. Run Times and Epochs of Learning by Scenario

	Social Welfare Maximizer		Monopolist	
	Run Time	Epochs	Run Time	Epochs
Sigma=0 MW	1 hr. 10 min.	300,000	1 hr. 10 min.	300,000
Sigma=150 MW	9 hr. 20 min.	2,400,000	5 hr. 50 min.	1,500,000
Sigma=300 MW	10 hr. 30 min.	2,700,000	7 hr. 47 min.	2,000,000

### 2.7 Conclusions

This essay demonstrates that RL is capable of modeling optimal investment behavior under uncertainty in an environment as complex as electrical power generation. Varying demands as well as multiple technologies from which firms may invest create

this complexity. This ability to model complex problems exists because the tabular Q-learning algorithm circumvents the curse of modeling by alleviating the need to explicitly define transition probabilities.

Investment problems that are ideal for solution using RL, compared with classical MDP solution methods, fall into two basic categories. The first class of problem, similar to the one addressed in this essay, uses complex algorithms to define state transitions and rewards. This class of problem is difficult to model using traditional MDP approaches because the formulation of transition probabilities may be nontrivial when dealing with multidimensional state representations.

The second class of problem that is ideal for the application of RL involves high dimension state representations such as investment in numerous technologies. RL has shown significant promise to solve those problems via use of Q-function approximators combined with its ability to ration computational time to high probability states. However, it is unlikely that optimal results could be attained via state space sweeps in problems with greater than three or four dimensions. Therefore, the application of Q-learning to high-dimension problems would most likely be limited to normativistic applications where sub-optimal solutions would still be quite useful.

This essay has highlighted the degree to which models of electricity generation investment can be biased if they treat uncertainty improperly. These results show significantly differing investment outcomes for varying levels of demand uncertainty. Both the failure to consider uncertainty and the overestimation of uncertainty can result

in poor predictions concerning actual investment outcomes. This issue is especially relevant when forecasting investment behavior in a restructured era in which “obligation to serve” agreements no longer exist. Forecasts of investment behavior from deterministic models may significantly overestimate actual investment levels and in turn fail to predict potential shortages in generation capacity during periods of peak demand. Additionally, if individual firms fail to incorporate uncertainty into their planning models, the market may provide investment that exceeds an efficient level.

Despite the strengths of RL, this essay also makes clear some of its drawbacks. First, although theorems exist that prove optimal RL convergence under certain conditions, these proofs usually guarantee optimality in infinite time. In practice, run times may be unreasonably long and highly sensitive to the model’s reward structure. For instance, in this application social welfare maximizing runs required significantly longer run times than profit maximizing runs. Additionally, as was reported previously, this application required a great deal of experimentation with the RL parameters (learning rate decay, action selection algorithm) to achieve near-optimal results with reasonable run-times. Algorithmic performance is highly sensitive to these parameters and ideal parameter selection is highly dependent upon the particular model. Therefore, there is no guarantee that the algorithmic modifications presented in this essay would be ideal for the application of RL to model investment behavior in other industries such as mining or petroleum. However, these parameters should serve as a good starting point for researchers who want to apply this research to other industries.

Future extensions to the model presented in this essay should include incorporating other technologies into the model. This could be accomplished in two ways. The first involves increasing the dimensionality of the technology state space. This would necessitate the use of a function approximator to estimate Q-values in lieu of the tabular approach utilized in this essay. A second approach would involve adding additional technologies to the initial capacity stock. This modification would affect reward calculations but would not increase the size of the state space as long as the agent could not invest in these additional technologies.

## Chapter 3

# THE EFFECT OF MARKET DESIGN ON ELECTRICITY GENERATION INVESTMENT UNDER DEMAND UNCERTAINTY

### 3.1 Background and Motivation

Most states in the United States are undergoing or considering restructuring that would establish some form of competition in the generation sector. One impetus for this change is the belief that electricity generation no longer possesses the subadditive cost properties of a natural monopoly due to technologically driven decreases in efficient plant sizes (Fox-Penner 1997). Therefore, restructuring may bring about efficiency gains which may lead to reduced customer prices and product innovation as was the case with the airline, telephone, natural gas, trucking, and railroad industries (Crandall and Ellig 1997).

Numerous studies have analyzed the short-run efficiency of restructured electricity markets (Borenstein and Bushnell 1998; Green and Newberry 1992; Wolak and Patrick 1996; Quick 2000). If generators can exert market power by varying the quantity or price of their bids, the spot price will exceed competitive levels and potentially offset any efficiency gains from restructuring. However, less attention has been paid to the long-run efficiency of restructuring—specifically, the area of investment in generation. This area is critical to understanding the implications of restructuring due

to the direct link between investment and reliability as well as the potential for investment-based efficiency gains. On the positive side, restructuring may bring about significant savings due to a more efficient investment level and a more efficient investment composition. However, policy makers who design markets must ensure that these gains are not made at the expense of reduced system reliability resulting from inadequate levels of generation.

Policy makers must establish “market rules” when setting up a restructured electricity market that may directly or indirectly affect the quantity and mix of generation investment that is provided by the market. This essay investigates how two of these market design decisions impact generation investment and electricity spot prices. Specifically, the essay examines capacity subsidies and spot market price caps. Several authors discuss the effects of capacity subsidies and price caps on generation investment and electricity price qualitatively, however, none show them quantitatively (Graves *et al.* 1998; Hirst, Kirby, and Hadley 1999; Singh and Jacobs 2000; Wolak *et al.* 1999).

Capacity subsidies, or reserve requirements, have been justified on the grounds that capacity possess the properties of a positive externality and therefore will be underprovided by the market. Price caps have been instituted in order to protect consumers from high prices that result from capacity scarcity or from strategic behavior by market participants. The two policies are related because both act to reduce spot prices during peak loads, which reduces the overall volatility of spot market prices. Capacity subsidies affect peak prices by increasing the total level of capacity that is

provided by the market; whereas, price caps affect prices directly by constraining the market-clearing price in the spot market. Because these policies act through different mechanisms, they produce differing effects on average price and investment. Capacity subsidies result in higher total electricity prices in addition to higher levels of investment. In contrast, price caps may result in higher or lower levels of investment, depending on the market structure. Additionally, price caps may require that loads be shed because they prevent price from rationing scarce supplies of energy.

The remainder of this essay is organized as follows: Section 3.2 discusses the restructured electricity environment, Section 3.3 discusses previous investment models that pertain to a restructured electricity market, Section 3.4 investigates the effect of capacity subsidies on investment, Section 3.5 analyzes the effect of price caps on investment, Section 3.6 examines the effect of the elasticity of demand on peak prices, and Section 3.7 provides policy suggestions and concluding remarks. Section 3.7 also provides a brief policy recommendation for the State of Colorado based upon these results.

### 3.2 The Restructured Environment

The majority of restructuring plans call for some kind of a spot market where generators sell to either distributors or customers directly. However, unlike other goods which can be bought and sold with little or no outside intervention, electricity markets must be closely controlled by a system operator who facilitates spot market operations

subject to physical system constraints (Fox-Penner 1997; Hogan 1998). These constraints are caused by the physical properties of electricity, most notably its inability to be inexpensively stored and the requirement that supply and demand must balance simultaneously. Another complicating factor in electricity markets is that loop flows may prevent power from flowing directly between a buyer and seller across transmission lines. Loop flows occur because electrical power obeys Kirchoff's law which will cause power to flow over the "path of least resistance" (Fox-Penner 1997, 27). To ensure the physical integrity of the system, electricity requires ancillary services which further complicate the design of electricity markets. Ancillary services include regulation, spinning reserves, non-spinning reserves, and replacement reserves. Ancillary services would be used, for example, if there were an unexpected supply disruption from a given plant. In this situation, spinning reserves could be brought online so that the supply of electricity remained unaffected (Fox-Penner 1997, 33).

One of the most popular forms for power markets is the POOLCO in which a system operator takes bids from various plants for the price at which they are willing to provide power over a set time period—usually an hour. The system operator also takes demand bids for the same time period. Next, the system operator determines a merit order dispatch in which firms are ranked by marginal cost, subject to system security constraints. Figure 13 illustrates that a merit order dispatch forms a stepped supply curve for energy in a given time period. Also, Figure 13 demonstrates that the marginal plant (#5) sets the spot market price in the period under consideration.

An alternate approach is a bilateral market in which buyers and sellers directly contract with each other for a given price and time period that both parties agree upon. For example, a generator could contract with a large industrial customer to provide power at a fixed price for a given time period. This sort of arrangement protects the industrial customer from price volatility and provides the generator with a certain revenue stream over the period of the contract.

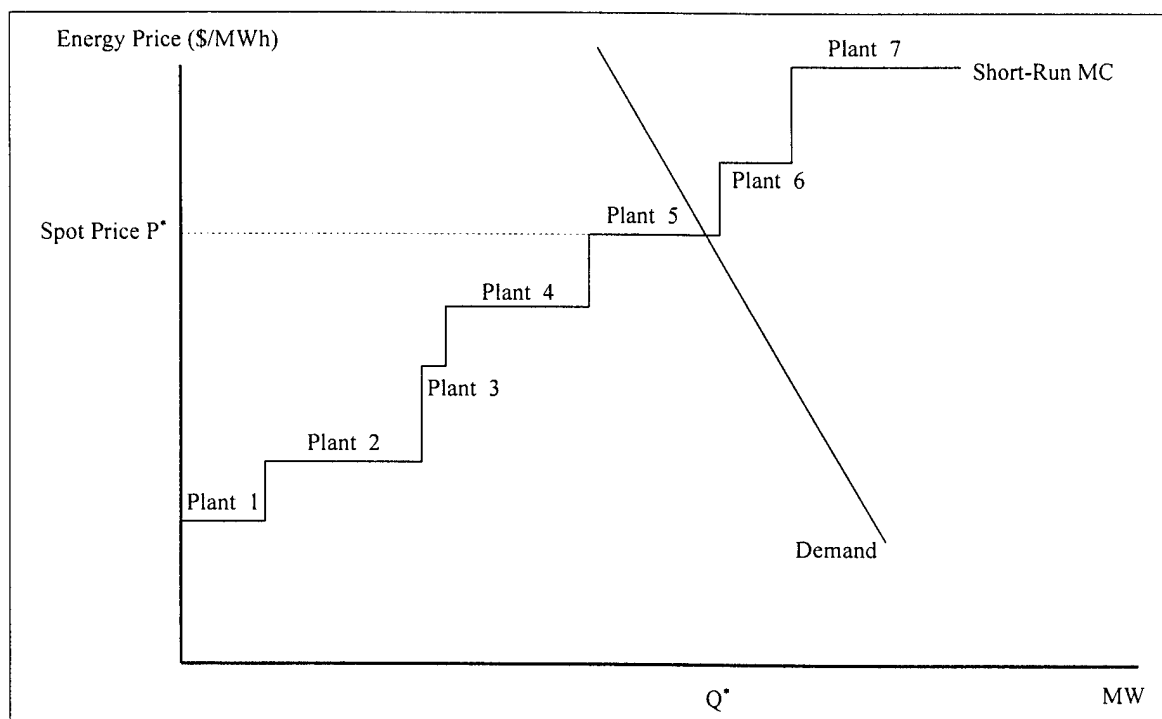


Figure 13. Short-Run Spot Market

As is the case with the POOLCO, a system operator must ensure that trades are feasible with respect to system security constraints. In practice, most markets involve a combination of POOLCO and bilateral designs (Fox-Penner 1997).

### 3.3 Models of Generation Investment under Restructuring

In order to investigate long-run effects of restructuring on investment, several studies have analyzed generation investment behavior in a restructured environment from both qualitative and quantitative perspectives. Orr (1988) uses an option-valuation approach to examine the effects of restructuring on capacity timing and technology choice for a monopolistic firm in both regulated and unregulated environments. He finds that restructuring will bring about the adoption of more fuel-efficient technologies sooner than remaining in a regulated environment. He also determines that the presence of demand uncertainty will bring about the more rapid adoption of newer technologies due to the added flexibility that they provide.

Fehr and Harbord (1997) consider the effects of oligopolistic energy markets on investment behavior. They determine that overall investment in “most reasonable cases” falls short of socially optimal levels. This decrease in investment results from the decrease in quantity produced by strategic firms in order to exert market power. (Fehr and Harbord 1997).

Hirst *et al.* (1999) use a hybrid optimization/simulation approach to determine the relationship between the reserve margin and total social costs while assuming perfectly inelastic short run demand. In order to ensure that energy markets clear when demand exceeds capacity, they implement an “unserved energy elasticity of demand” (UEED). This is defined as a price elasticity of demand that is activated only when demand exceeds capacity. Their results show that, for a UEED of 0.05, reserve margins from 2 to

7 percent minimize total social costs with margins outside this range leading to significant increases in total social costs. Their social cost calculations include the price of energy, any necessary capacity payments, and the social costs of not meeting demand due to insufficient capacity. These social costs are assessed when demand exceeds capacity and the UEED is utilized (Hirst *et al.* 1999). Because Hirst *et al.* do not consider ancillary services, they suggest that these estimated reserve margins should be increased by 5 percent to determine actual reserve requirements.

### 3.4 The Effect of Capacity Subsidies on Generation Investment

Several restructured electricity markets have elected to institute reserve requirements or capacity subsidies whereas some have not. Section 3.4.1 discusses the motivations behind the different approaches and provides descriptions of several actual market designs, Section 3.4.2 presents a reinforcement learning (RL) model of investment that quantifies the effects of capacity subsidies on generation investment and electricity spot price, and Section 3.4.3 summarizes the results from the capacity subsidy model.

#### 3.4.1 Background on Capacity Subsidies and Reserve Requirements

A reliable electricity system can be defined as one “that allows for few involuntary interruptions of service to customers” (DOE 1998). This encompassing definition can be broken down into two components, namely adequacy and security.

Adequacy, which is a long-run planning concept, refers to maintaining an adequate quantity of generation to meet supply. This contrasts with security, which is a short run planning concept that refers to the ability to respond to short-run disturbances in electrical supply (DOE 1998; Hirst, Kirby, and Hadley 1999). Of these concepts, adequacy is the focus of this essay.

Under the regulated system, generation adequacy was assured through mandated reserve requirements that were established by regulators. This system paid generators separately for capacity and energy to ensure that all fixed costs could be recouped, especially on peaking plants that were used infrequently (Graves *et al.* 1998).

Under restructured systems, several market designs have been implemented in order to ensure generation adequacy. These designs include a strict reliance on markets to provide for a sufficient level of generation investment or direct intervention through either a capacity subsidy or a mandatory reserve requirement. The market-based design relies solely on price signals to motivate investment. This approach is often referred to as an “energy-only” system because energy is the only traded commodity. Capacity subsidies encourage generation investment by subsidizing firms directly for their capacity regardless of its dispatch status. Similarly, reserve requirements mandate that all market participants share the responsibility for providing for excess reserves. This requirement is usually enforced through fines on firms that do not meet their system operator-dictated capacity obligations. Markets with reserve requirements often establish separate capacity

trading markets so that firms may meet capacity requirements through investment or through the purchase of tradable capacity credits.

Some of the impetus for keeping capacity payments or reserve margins under restructuring may be attributed to path dependence from the regulated era. Additionally, many load serving entities (LSEs) favor keeping capacity payments in place because they provide a certain revenue stream for any new generation investment (Singh and Jacobs 2000).

One argument for a capacity subsidy or reserve requirement is that the inability of demanders to react to price in real time may prevent the market from clearing when demand exceeds capacity. Therefore, it is necessary for a system to possess sufficient reserves to meet peak loads. This argument is illustrated in Figure 14, which shows a perfectly inelastic demand ( $Q^*$ ) that exceeds total capacity ( $Q_c$ ). No equilibrium market-clearing price exists and the ISO must intervene by shedding load. When the ISO sheds load in this manner, it is likely that it will not be able to identify those customers who are most willing to curtail their load in return for some form of compensation. Therefore, it is likely that the allocation of scarce capacity will be inefficient. In contrast, a system that promotes a demand-side response to price by allowing customers to self-select their level of reliability will efficiently ration scarce capacity levels through increases in price. This sort of system also requires less information on the part of the regulator than a market with a perfectly inelastic demand.

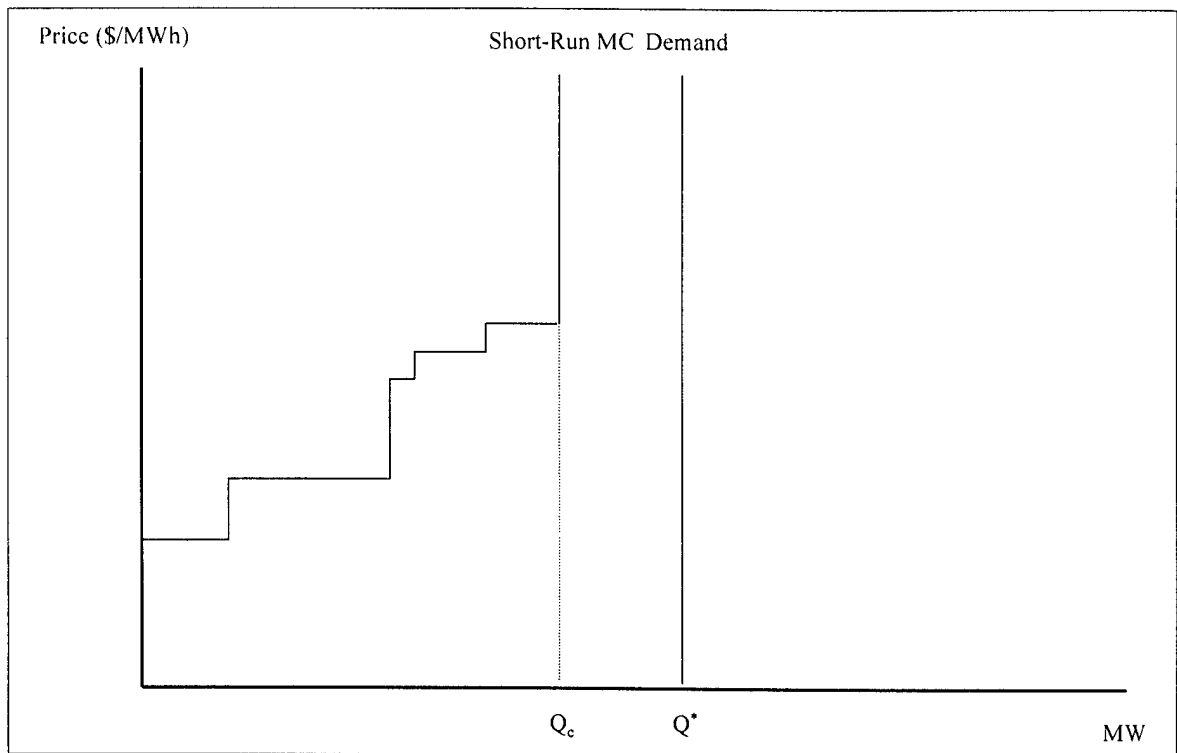


Figure 14. No Market Clearing Price with Inelastic Demand

Another argument for either a capacity subsidy or a reserve requirement is the belief that capacity possesses the properties of a positive externality and therefore will be underprovided by the market. This contention is based upon the idea that an individual firm's excess capacity benefits all market participants because the ISO may have to override economic relationships between market participants in order to maintain the physical security of the system. The probability of the ISO intervening in this manner decreases as the excess capacity in the system increases (Jaffe and Felder 1996; Ruff 1999).

This argument has been criticized on the grounds that individual firms do realize benefits from their excess capacity as long as an efficient ancillary services market exists that compensates firms for their excess capacity based upon its value to the system (Borenstein 1999a). Therefore, opponents of the positive externality argument suggest that capacity subsidies or reserve requirements will lead to inefficient investment both in terms of technology composition and overall investment level. This results from the removal of price as a signal for firms to increase their capacity. Rather, firms will invest to meet mandatory reserve requirements in the most cost-effective manner possible (Graves *et al.* 1998).

Another argument against capacity subsidies or reserve requirements is that, even if the positive externality argument is correct, these approaches assume that a regulator knows with certainty the efficient capacity subsidy or reserve requirement. This requires knowledge of the marginal cost of adding new capacity as well as the marginal social benefit function. While the marginal costs of adding capacity are reasonable to estimate, it is difficult to determine the marginal social benefit function that results from adding excess capacity. If this estimate is incorrect, then welfare losses from these policies could greatly exceed the welfare losses from imposing no subsidy or reserve requirement (Graves *et al.* 1998; Jaffe and Felder 1996).

This principal is illustrated in Figure 15, which graphs the marginal cost to society of adding reserves as well as the marginal social benefits of excess reserves. For this example it is assumed that capacity is a positive externality and that no ancillary services

market exists (Jaffe and Felder 1996).  $R_c$  represents the quantity of reserves that would be supplied with no capacity subsidy or reserve requirement. The marginal cost to society increases from zero, because initial investments in reserves will possess some private benefits because these reserves will most likely be dispatched occasionally. This marginal social cost function becomes constant once reserves are being added that will never be dispatched. In this situation, the entire cost of reserves must be subsidized in order for them to be built. The marginal social benefit function takes the shape of a negative exponential distribution because the probability of an outage is a negative exponential function of available reserves (Stoll 1989, 331).

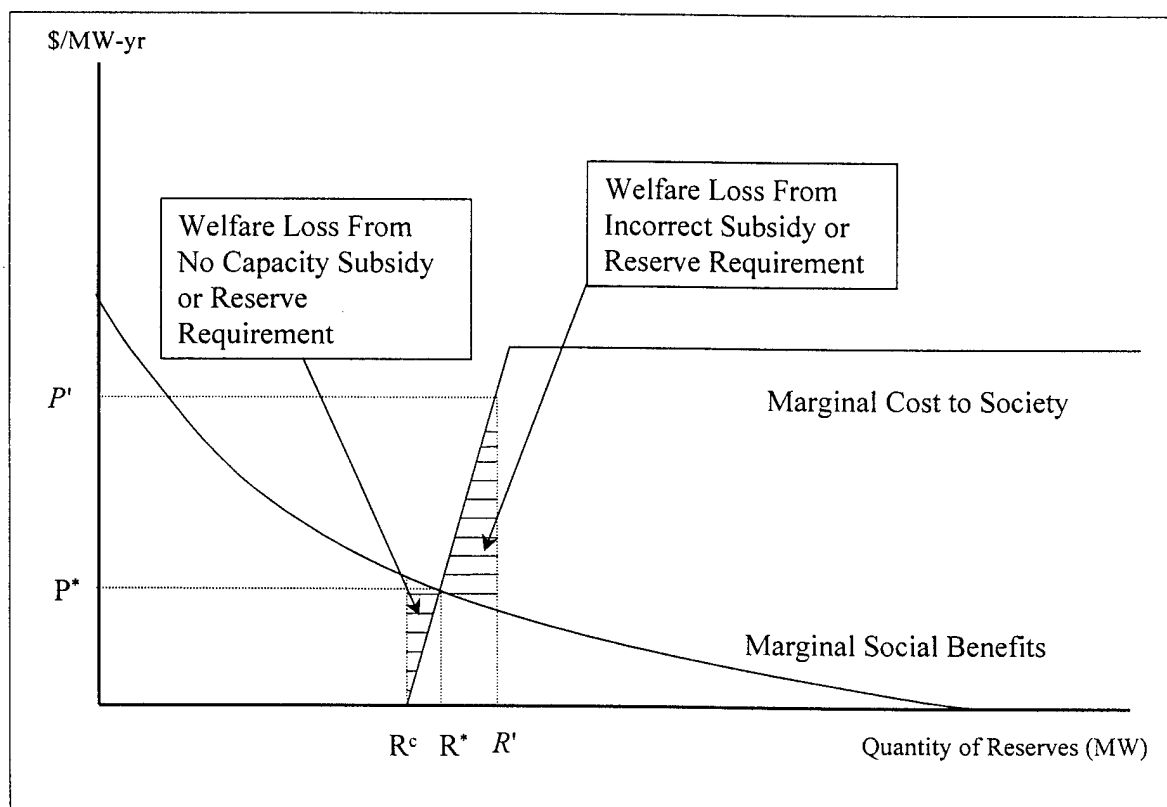


Figure 15. Welfare Loss from an Inefficient Reserve Requirement or Capacity Subsidy

This figure illustrates that either an optimally set subsidy  $P^*$  or reserve requirement  $R^*$  will produce an efficient outcome. Additionally, it illustrates that the welfare loss from an incorrect subsidy  $P'$  or reserve requirement  $R'$  may produce a welfare loss that exceeds the loss from no policy at all.

Finally, Singh and Jacobs (2000) suggest that mandatory reserve requirements may actually do little to augment generation adequacy because excess capacity that is built to meet a given standard is often bid elsewhere during peak loads. They note that this occurs when dispatch systems that do not implement mandatory reserve requirements border reserve requirement-based systems. Those systems without reserve requirements will likely have higher spot prices than regions with reserve requirements during peak loads due to their lower capacity obligations. In this situation, firms in the region with the reserve requirement may sell their power outside the region to take advantage of the higher outside price. This scenario highlights why reserve requirements may have been more appropriate under the regulated system, where each franchised monopoly was “obligated” to serve its load, than a restructured system.

These contrasting arguments have lead different regions in the United States to implement systems based on either reserve requirements or energy-only markets. No direct capacity subsidies have been enacted in the United States, however, countries such as the United Kingdom, Spain and Argentina utilize capacity subsidies. Singh and Jacobs ( 2000) note that in the United States “many capacity requirements often reduce to capacity payments.” This occurs in some systems because fines collected due to

noncompliance with capacity obligations may be distributed to all firms that have capacity in excess of their obligations thus resulting in a subsidy to those firms.

The next section of this essay describes the market designs in California and the Pennsylvania-New Jersey-Maryland Interconnection (PJM) as examples of energy-only and reserve requirement-based systems. The California market does contain a reserve market for ancillary services; however, since this market's purpose is the maintenance of security as opposed to adequacy, the California market will still be referred to as an energy only market.

#### 3.4.1.1 California Market Design (Energy-Only Market)

California has implemented one of the most progressive market designs in the nation. Its restructured system involves two independent institutions, the Power Exchange (PX) and the Independent System Operator (ISO). The PX runs a day-ahead spot market where supply and demand bids are accepted for each hour of the subsequent day. Sellers to the PX include independently owned utilities (IOUs) and distribution companies. Demanders include power marketers and industrial customers. The ISO has two primary responsibilities. First, it ensures that supply and demand bids by the PX, as well as bilateral contracts that were made outside the PX, are feasible given system security constraints. Secondly, the ISO runs an ancillary services market for "real-time" generation in order to keep supply and demand in balance. Firms may bid any nondispatched capacity into the ancillary services market. Ancillary service costs are

passed onto consumers (Borenstein and Bushnell 1998). Additionally, some plants have been designated as must-take power plants and are exempt from bidding into the PX. Instead, the ISO must accept power from these plants at prearranged prices. Examples of must take plants include certain nuclear and hydro plants (Graves *et al.* 1998).

In the California system, neither the ISO nor the PX mandate any level of required reserves. Similarly, no separate capacity trading market or capacity subsidy exists for owners of generation capacity (Hirst *et al.* 1999). The success of California's energy-only system is still being debated since the restructured marketplace has only been operating since April of 1998. The California ISO (1998) defends its system by highlighting that a sufficient quantity of capacity additions are planned to meet forecast demands. However, some argue that California's system will lead to adequacy problems as a result of its strict reliance on markets (Conkling 1998; Michaels 1997). Graves *et al.* refute that claim and state that "regulators are pursuing restructuring precisely because past capacity decisions based upon uniform reliability criteria have not produced an economical supply mix" (Graves *et al.* 1998).

#### 3.4.1.2 PJM Market Design (Reserve Requirement Market)

The PJM design is representative of reserve requirement-based power markets in the northeastern United States. The ISO New York and ISO New England (NEPOOL) are similar in structure to the PJM. The PJM is the largest centrally dispatched electric control area in North America and the third largest in the world. It includes sections of

Pennsylvania, New Jersey, Maryland, Delaware, Virginia, and the District of Columbia (PJM 1999). Unlike the California system that puts no explicit requirement on capacity, the PJM system requires that all LSEs provide a fraction of an aggregate reserve requirement that the PJM Reliability Committee deems necessary (PJM 1998a). This reserve requirement is based upon the amount of excess capacity that is needed to “ensure a sufficient amount of capacity to meet the forecast load plus reserves adequate to provide for the unavailability of capacity resources, load forecasting uncertainty, and planned maintenance outages” (PJM 1998a). Current PJM reserve requirements mandate reserves of at least 19.5 percent (Bhavaraju 1999).

Additionally, the PJM Office of the Interconnection operates voluntary monthly and mandatory day-ahead capacity markets where capacity credits can be sold or bought so that individual firms can buy capacity credits if that option is cheaper than investing in excess capacity themselves (PJM 1998b). The system of tradable capacity permits is analogous to a tradable emissions permit system (Jaffe and Felder 1996). Firms may voluntarily provide capacity supply and demand bids to the monthly market. In contrast, participation in the day-ahead market is mandatory for all firms with capacity levels above or below their capacity obligation. Firms with excess capacity that do not submit bids will have bids submitted for them at \$0/MW-day. Similarly, firms with deficient capacity positions that do not bid will have bids placed for them at the Capacity Deficiency Rate (CDR) of \$158/MW-day (PJM 2000). The CDR is also assessed to all firms that do not meet their specified capacity obligations either through capacity

investments or the purchase of credits. All collected CDR payments are distributed to firms with surplus levels of capacity based on the amount of excess capacity that they hold, thus acting as a capacity subsidy (Singh and Jacobs 2000).

This market design favors adequacy assurance at the expense of potential reductions in cost. Henney (1998) claims that the PJM reserve requirement obscures price signals to investors and creates barriers to entry for potential entrants by increasing the complexity of the system. Another criticism concerns the assumption, when calculating reserve requirements, that outages can be described using a Poisson distribution. This distribution assumes that the probability of a plant outage is independent of the time period under consideration. The Poisson distribution may be inappropriate because the high electricity spot prices during peak load periods may cause these periods to experience fewer outages than non-peak periods (Graves *et al.* 1998).

Singh and Jacobs (2000) site the PJM as an example of a market where a reserve requirement does little to improve adequacy because the neighboring East Central Area Reliability Council (ECAR) does not possess a reserve requirement. They show that on hot summer days, firms in the PJM “delist” themselves as available and then sell their power to the neighboring ECAR. Firms exhibit this behavior despite the fact that they may be liable for CDR payments due to their failure to meet their capacity obligations.

### 3.4.2 Model of Capacity Subsidies

The reinforcement learning model that is presented in Section 2.4 of the first essay is modified to determine the effect of capacity subsidies on generation investment and electricity prices. As was the case in Chapter 2, this model utilizes an iso-elastic demand curve with an elasticity of 0.1. All other modeling assumptions presented in Section 2.4.1 are applicable to this model.

Capacity subsidies are implemented in the RL model by first calculating a capacity subsidy and then adding this subsidy to the previous calculated reward that is discussed in Section 2.4.2.5. A firm's capacity subsidy is equal to the product of its total capacity and the per-MW capacity payment. It is assumed that this subsidy is financed by consumers who pay a per-MWh capacity charge that is added to the per-MWh electricity wholesale price after one year of dispatch is complete. For the purposes of this essay, the average electricity spot price plus this capacity charge is defined as the total price of electricity. This total electricity price is not the actual consumer price because transmission, distribution, and ancillary service charges are not included.

The model considers subsidy levels ranging from \$0/MW-yr to \$60,000/MW-yr in increments of \$20,000/MW-yr for the social welfare maximizing agent. These levels ensure that the capacity payments are comparable with observed capacity prices in markets where capacity is traded. Several representative values for capacity prices are listed in Table 5.

Table 5. Observed Capacity Values

Market	Capacity Price (\$/MW-yr)
PJM Monthly Capacity Market for Jul 1999 (12 month high)	43,800
PJM Monthly Capacity Market for March 2000 (12 month low)	1,825
NEPOOL Capacity Price for April 1999 (high from Apr 98 – Jan 00)	14,916
Proposed Colorado Capacity Payment (Stone and Webster 1999)	11,110
PJM Capacity Deficiency Rate (Penalty imposed on PJM firms not meeting capacity requirement)	57,998

The first three rows show equilibrium market prices from the monthly PJM and NEPOOL capacity markets. The fourth row shows the value of a proposed capacity subsidy that was utilized in a recent study of electricity restructuring in the state of Colorado (Stone and Webster 1998). Finally, the last row of the table shows the PJM CDR rate in \$/MW-yr.

Capacity subsidies are modeled rather than reserve requirements because it is difficult to apply reserve requirements to demand curves with nonzero elasticities. Reserve requirement calculations traditionally assume that demand is perfectly inelastic and reserve requirements  $R$  equal:

$$R = \frac{K - D}{K}, \quad (3.1)$$

where,  $K$  represents capacity and  $D$  represents demand. Under these assumptions, any addition to capacity will lead to a direct increase in reserves assuming that capacity exceeds demand. However, when demand curves with nonzero elasticities are used,

calculation of reserves becomes more complex because the quantity of energy demanded is a function of the equilibrium price. Therefore, additions to capacity do not always lead to increases in reserves. This principal is illustrated in Figure 16 which shows an initial capacity level  $K_1$  along with two augmented capacity levels  $K_2$  and  $K_3$ . The increase in capacity from  $K_1$  to  $K_2$  with a marginal cost of  $P_2$  has no impact on reserves because the quantity of energy demanded increases when price falls from  $P_1$  to  $P_2$ . Reserves do not increase until capacity is increased above  $K_2$ .

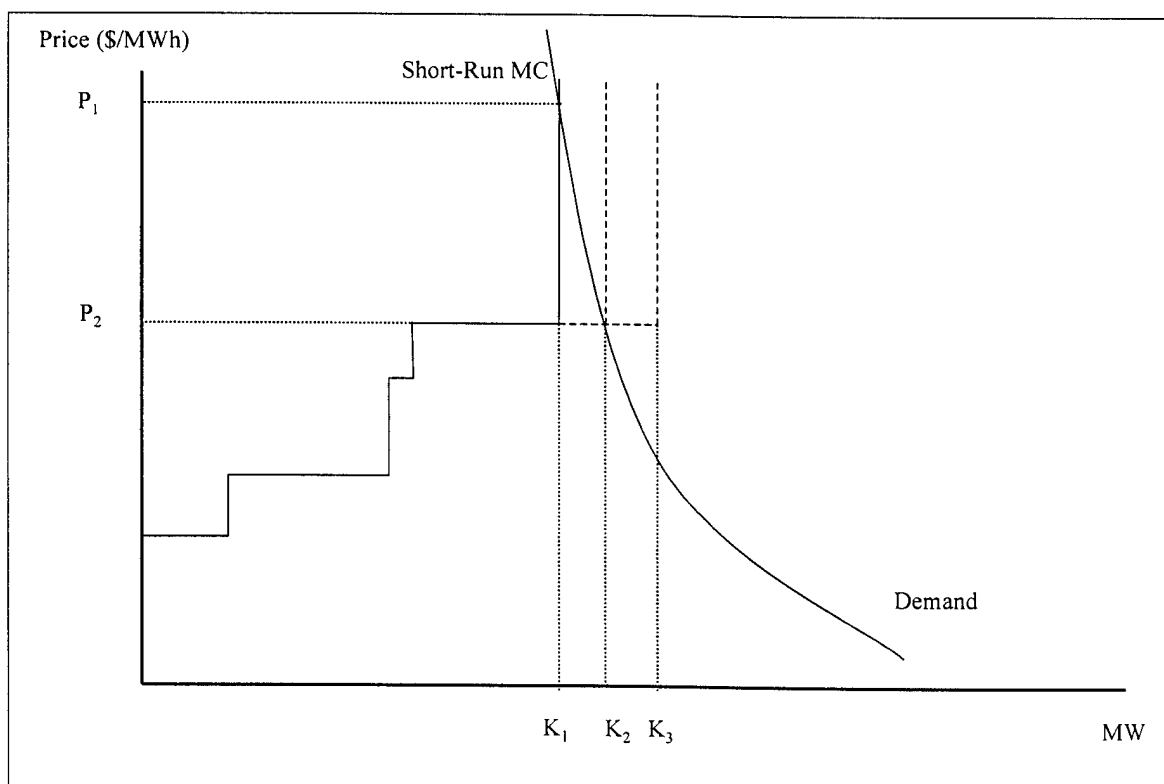


Figure 16. Calculation of Reserve Requirement Using Price Responsive Demand Curve

Therefore, setting a reserve margin may result in a very large increase in overall capacity. In contrast, capacity subsidies result in levels of investment that are roughly proportional to the level of the subsidy as is shown in Figure 17 of the next section.

The short-run efficiency of a correctly set capacity subsidy is equivalent to that of a capacity standard. However, capacity subsidies and capacity standards may have different distributional effects and thus different long-run outcomes as firms exit or enter the industry (Jaffe and Felder 1996; Weitzman 1974). Since demand uncertainty exists, the direct equivalence between standards and subsidies can not be assumed, even in the short-run, because the value of adding reserves varies based upon the most recent realization of demand. If an abnormally large increase in demand occurs, the social benefit of adding capacity is greater than the social benefit of adding capacity following a decrease in demand. However, despite this lack of a direct equivalence between the two mechanisms, investment behavior under capacity requirements should be similar to that under capacity subsidies. Also, since reserve requirements are a special case of capacity standards that mandate an excess quantity of capacity at peak loads, investment under a reserve requirement should be similar to investment under a capacity subsidy.

### 3.4.3 Capacity Subsidy Results

Mean capacity levels for varying subsidy levels are graphed in Figure 17. As expected, higher capacity subsidies produce higher capacity levels with the highest subsidy level of

\$60,000/MW-yr producing more than twice the mean rate of investment compared with the non-subsidized level.

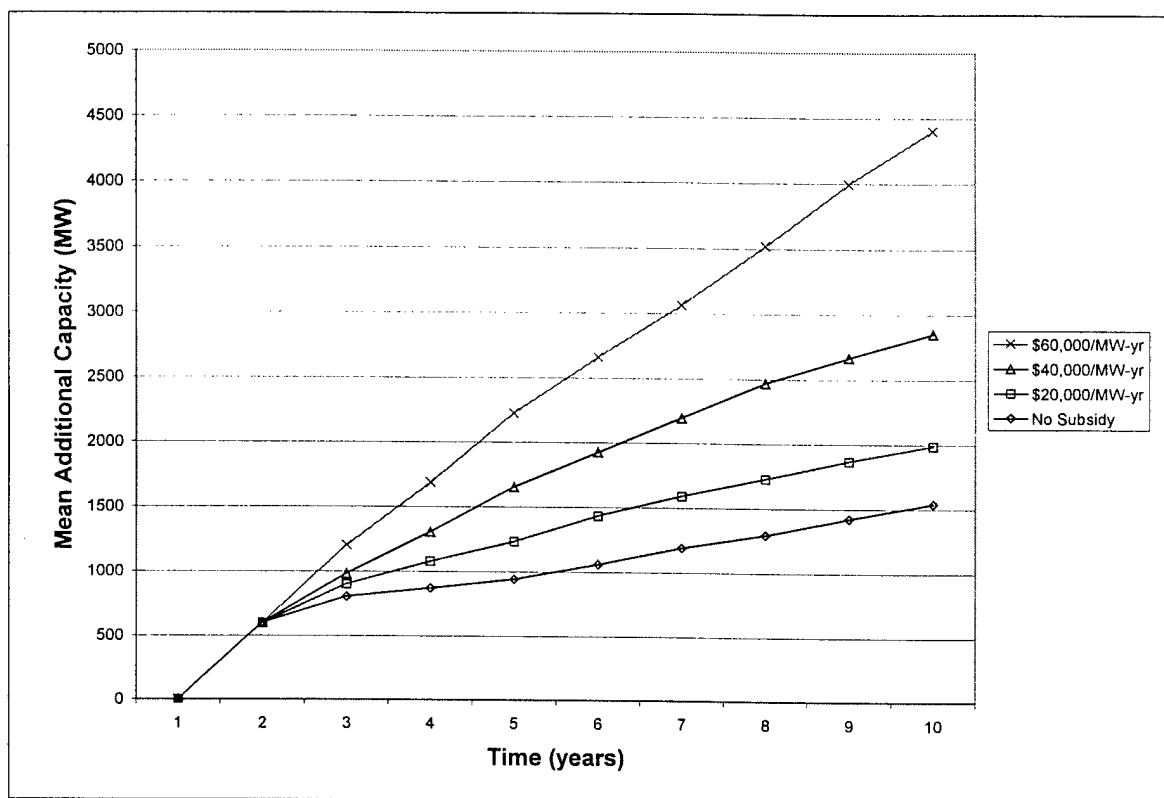


Figure 17. The Effect of Capacity Subsidies

These higher levels of capacity act to reduce peak prices by shifting the vertical portion of the supply curve outward, as is shown in Figure 16. This reduces the scarcity premium observed during peak demand periods, which in turn reduces peak wholesale prices. This effect is illustrated in Figure 18 which plots 10 years of mean peak prices for each subsidy level.

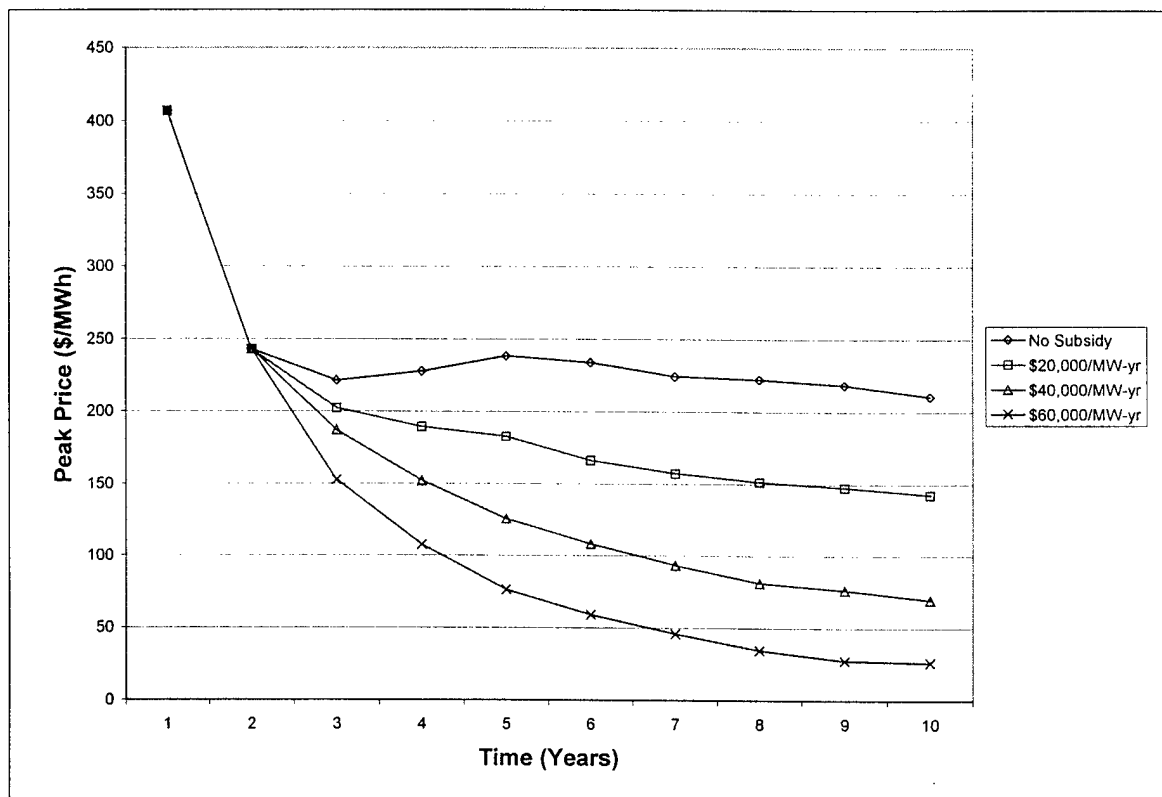


Figure 18. Effect of Capacity Subsidies on Peak Price

Since peak and near-peak prices are reduced for higher capacity subsidy levels, higher subsidies also result in a reduction in overall spot market volatility and mean spot market price. The standard deviation and mean of the spot market price for year 10 are plotted in Figure 19. Similar results exist for the other years.

These reductions in mean price and volatility do not come without a cost. A Per MWh capacity charge can be computed by dividing total capacity payments for a given year by the total number of MWh that were dispatched in that year.

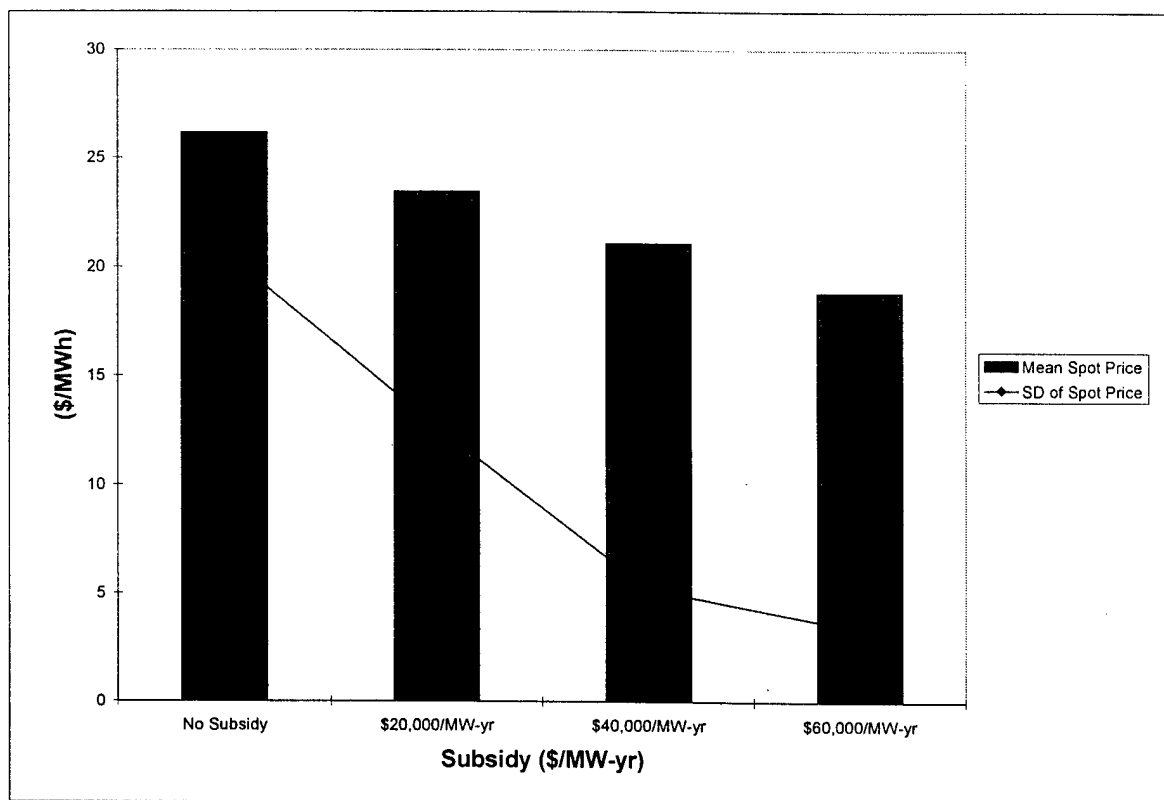


Figure 19. Effect of Capacity Subsidies on Spot Price

When this charge is added to the mean price, a total price for electricity in \$/MWh can be computed. This total price covers both energy and capacity subsidy costs. This total price of electricity increases with increasing capacity subsidy levels. Singh and Jacobs (2000) suggest an alternate means for interpreting these results which equates capacity subsidies or reserve requirements with a call option on electricity. Under this interpretation, the increase in total electricity price associated with a capacity subsidy or a reserve requirement is analogous to the price of the option. This option protects consumers from upward price movements and is “exercised” when excess capacity, originating from either capacity subsidies or reserve requirements, is utilized during peak

demand periods. Mean spot market price and capacity charges are plotted in Figure 20 for each of the capacity subsidy levels. Note that total electricity price increases as the subsidy increases.

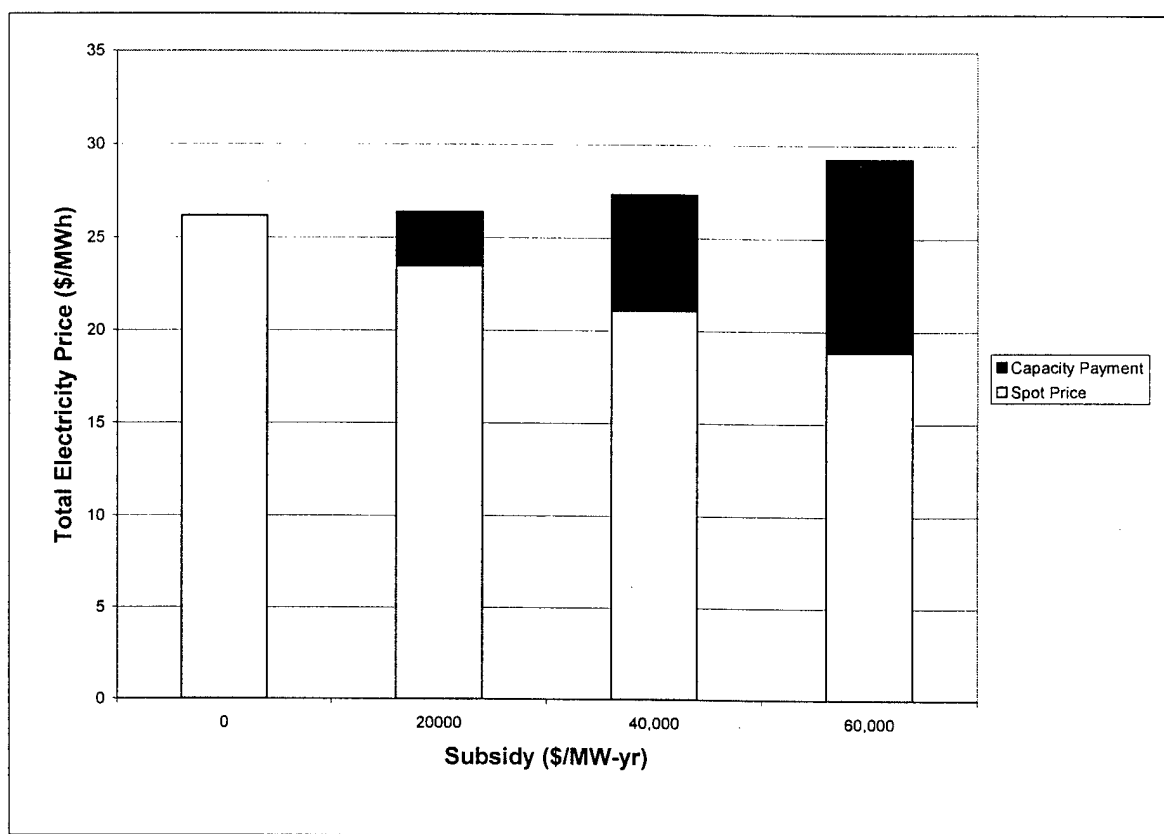


Figure 20. Effect of Capacity Subsidies on Total Price

In addition to altering the overall level of investment, capacity subsidies affect the composition of investment by increasing the overall percentage of investment in peaking generation (combustion turbine). This effect is illustrated in Figure 21 which plots the

mean percentage of the total additional capacity that is comprised of combustion turbine (CT) technology for years 2 through 10.

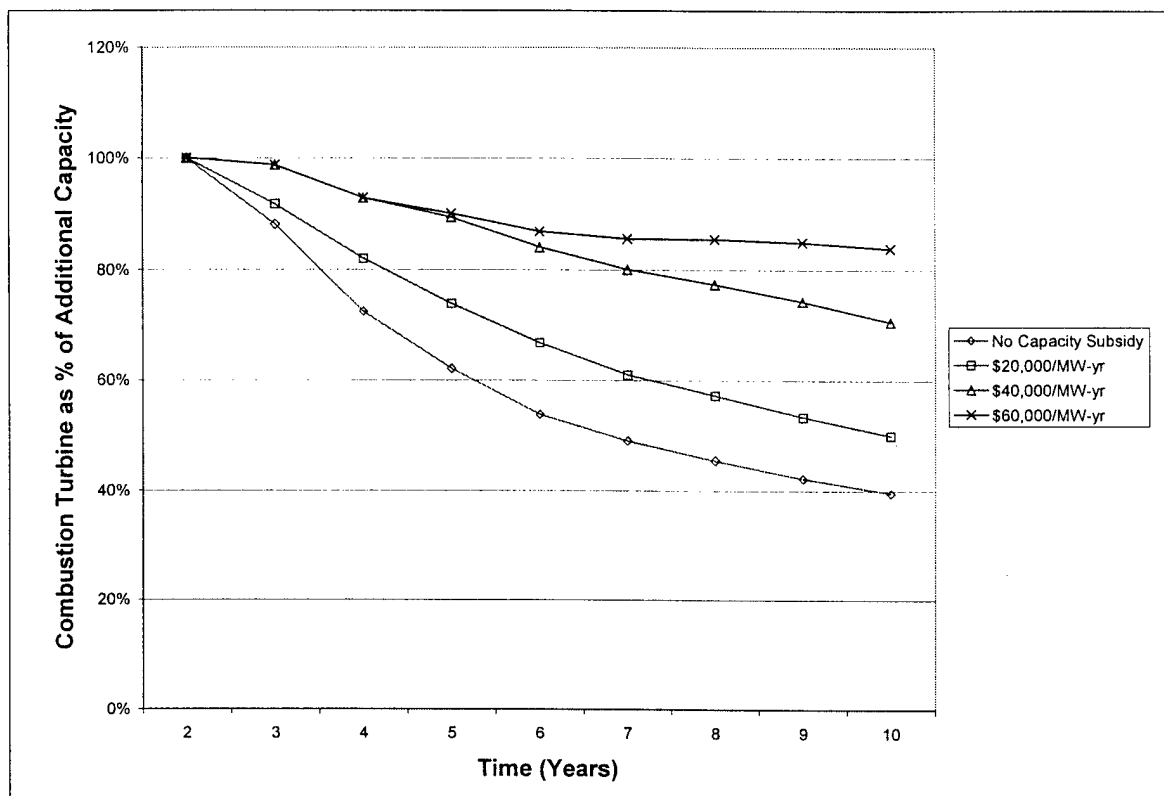


Figure 21. Effect of Capacity Subsidies on Investment Composition

Year 1 is not plotted because no additional generation exists in year 1 as a result of the one year delay between the investment decision and capacity becoming operational.

CT investment increases at higher capacity subsidy levels because this technology is more cost-effective for meeting peak loads than combined cycle (CC) generation because of its low up front investment and per-period fixed costs. CT generation is also a more cost-effective option for capacity that is seldom dispatched and is invested in for the sole purpose of receiving capacity payments.

### 3.5 The Effect of Price Caps on Generation Investment

Another relevant market design issue is whether or not to impose price caps on the spot market. A secondary decision involves setting the level of the cap if a cap is deemed necessary. Some markets such as California and the PJM have implemented price caps while others such as NEPOOL have not. Section 3.5.1 provides background on price caps, Section 3.5.2 presents a RL-based model of investment that incorporates price caps, and Section 3.5.3 summarizes results from this model.

#### 3.5.1 Why Price Caps May be Implemented

In a perfectly competitive market, electricity prices will be at their highest level during periods of peak demand for two reasons. The first is the basic economic principal that the plants with the highest marginal costs will be needed during these periods. The second reason, which further inflates prices during these peak periods, is that when capacity is scarce, the equilibrium price will rise above the marginal cost of the highest marginal cost plant so that the market will clear (Borenstein 1999b; Graves *et al.* 1998). Therefore, a capacity premium or scarcity premium emerges in these periods reflecting the scarcity of capacity. This capacity premium is illustrated in Figure 22. Therefore, supramarginal bids during peak demand periods are not necessarily signs of market power (Borenstein 1999b; Graves *et al.* 1998). As shown in Figure 14, this method of

rationing scarce capacity cannot take place if demand is perfectly inelastic and customers have no demand response to increased prices.

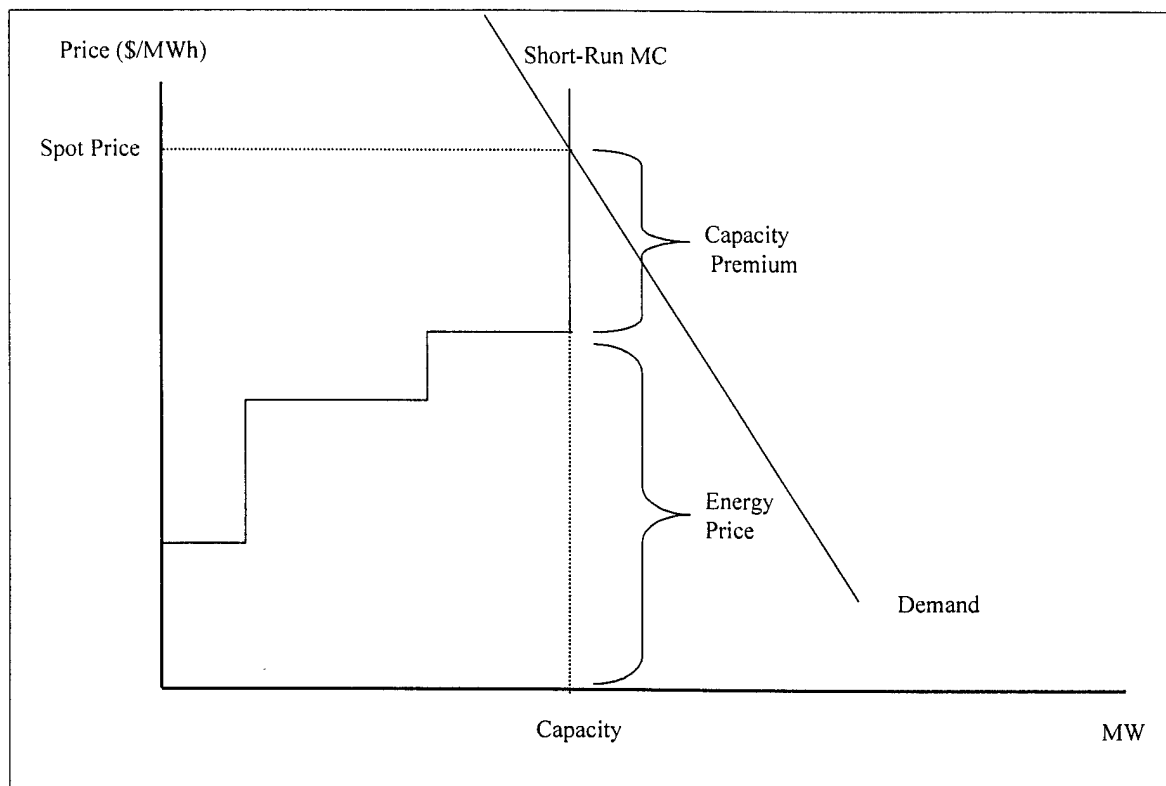


Figure 22. Capacity Premium during Periods of Peak Demand

Prices may also rise above competitive levels due to strategic behavior by market participants. Market power at times of peak load combined with an inelastic demand response can lead to extremely high spot prices during peak loads. This problem can be exacerbated when firms strategically congest transmission lines in order to increase their market power (Quick 2000). In the ancillary services markets, few supply bids and a lack of a demand-side response to price can further exacerbate this problem. In fact, under certain situations, firms can receive nearly any price they bid. An example of this

occurred on July 13, 1998, when the California regulation ancillary services market cleared at a price of \$9999 per MWh—the highest price that the ISO would allow. On this occasion especially high demand coupled with several plant failures created a situation where the ISO had no choice but to accept any bid price (Wolak *et al.* 1999).

These high prices, whether they are efficient competitive prices or results of strategic behavior, have motivated several markets to institute price caps in both energy spot markets and ancillary services markets. Some have argued that the public must be protected from high prices, even if they accurately reflect the true scarcity of energy and capacity (Graves *et al.* 1998; Hirst, Kirby, and Hadley 1999). Others have argued that price caps only have a role for mitigating the high prices that result from strategic behavior. This justification has been used to explain California's use of a \$750/MWh price cap for both its real-time energy and ancillary services markets. Some proponents of these price caps have argued that they are merely necessary transitional measures that will not be needed once customers are exposed to real-time electricity prices (Wolak *et al.* 1999).

### 3.5.2 Modeling the Effects of Price Caps

The long-run effects of price caps are analyzed using an enhanced version of the RL-based model of electricity generation investment that is presented in Section 2.4 of the first essay. The enhanced model imposes a price cap on the energy spot market.

Other than the imposed price cap, this model is identical to the model described in the first essay with respect to its assumptions and basic structure.

Price caps are modeled by forcing spot price to equal some specified level  $P_c$  if the market clearing price is greater than this value for every load duration curve segment. Since all price cap levels that are considered fall above the marginal cost of both CC and CT technologies, the price cap will only be binding in the vertical portion of the supply curve. Therefore, when the price cap is binding, suppliers will be producing at capacity and the quantity of energy that is demanded will exceed this level. The ISO will be forced to shed load to prevent system failure rather than allowing the market to force load reductions through higher prices. This scenario is illustrated in Figure 23. In this figure,  $P^s$  represents price without the price cap. Quantity  $Q^{pc}$  represents the quantity that would be demanded under the price cap if not for the capacity constraint. Since actual output under the cap must be set equal to capacity by shedding load, the quantity of load that must be shed by the ISO is equal to  $Q^{pc}$ -capacity.

This model is run for price caps ranging from \$30/MWh to \$200/MWh in \$10/MWh increments and from \$200/MWh to \$800/MWh in \$100/MWh increments for both monopolistic and social welfare maximizing perspectives. Increments are smaller for the lower price cap range (\$30/MWh-\$200/MWh) because price caps in this range have significant effects that differ based upon minute changes in the cap level. This contrasts with effects of cap movement in the higher range (\$200/MWh-\$800/MWh) where little or no changes are observed.

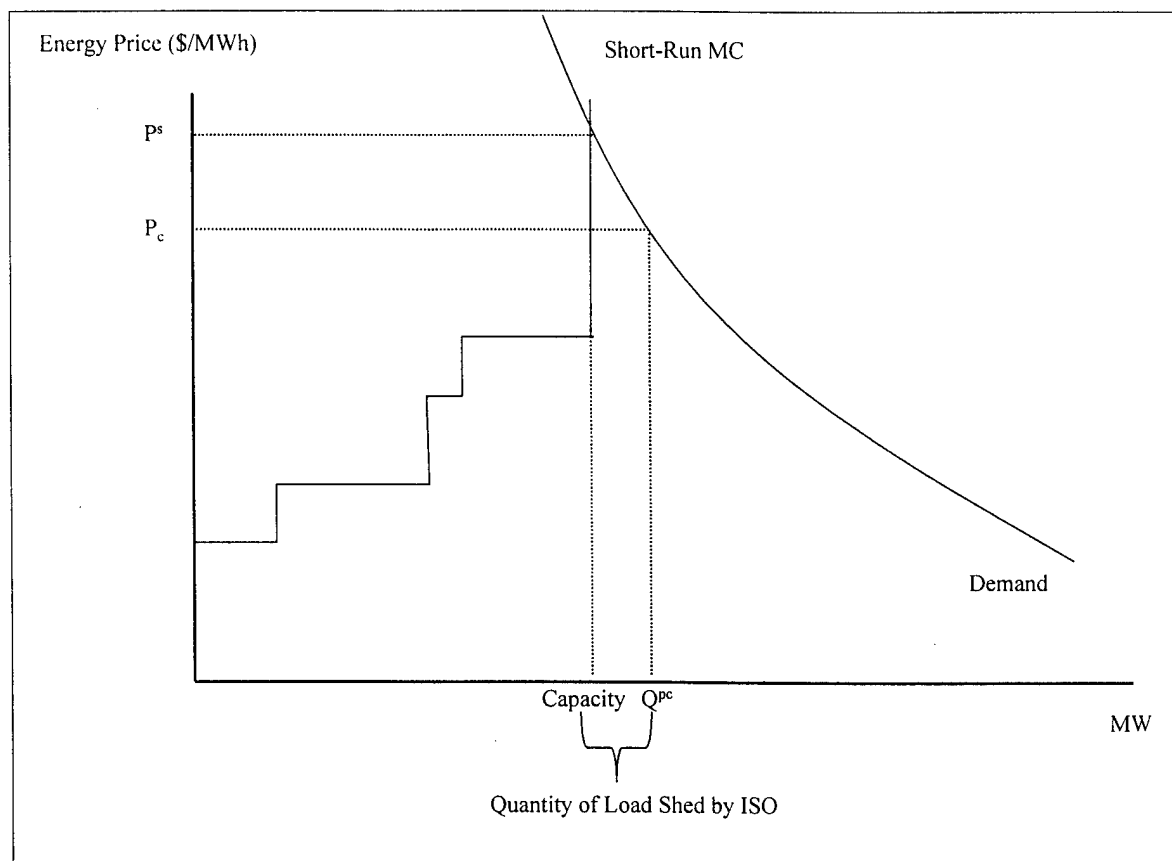


Figure 23. The Effects of a Price Cap

### 3.5.3 The Effect of Price Caps—Results

Mean monopolist capacity in year 10 for varying price cap levels and varying initial capacity levels is illustrated in Figure 24. This figure also plots mean spot price for each cap level on the right-hand axis. Simulation results for these price cap levels were initiated from 8,000 MW in addition to the baseline level of 10,000 MW in order to illustrate the differences between price cap levels. For the baseline initial capacity level of 10,000 MW, price caps above \$300/MWh did not show any additional investment by

year 10. Note that the effect of the price cap on investment is bi-modal. Three separate effects can explain the shape of this graph.

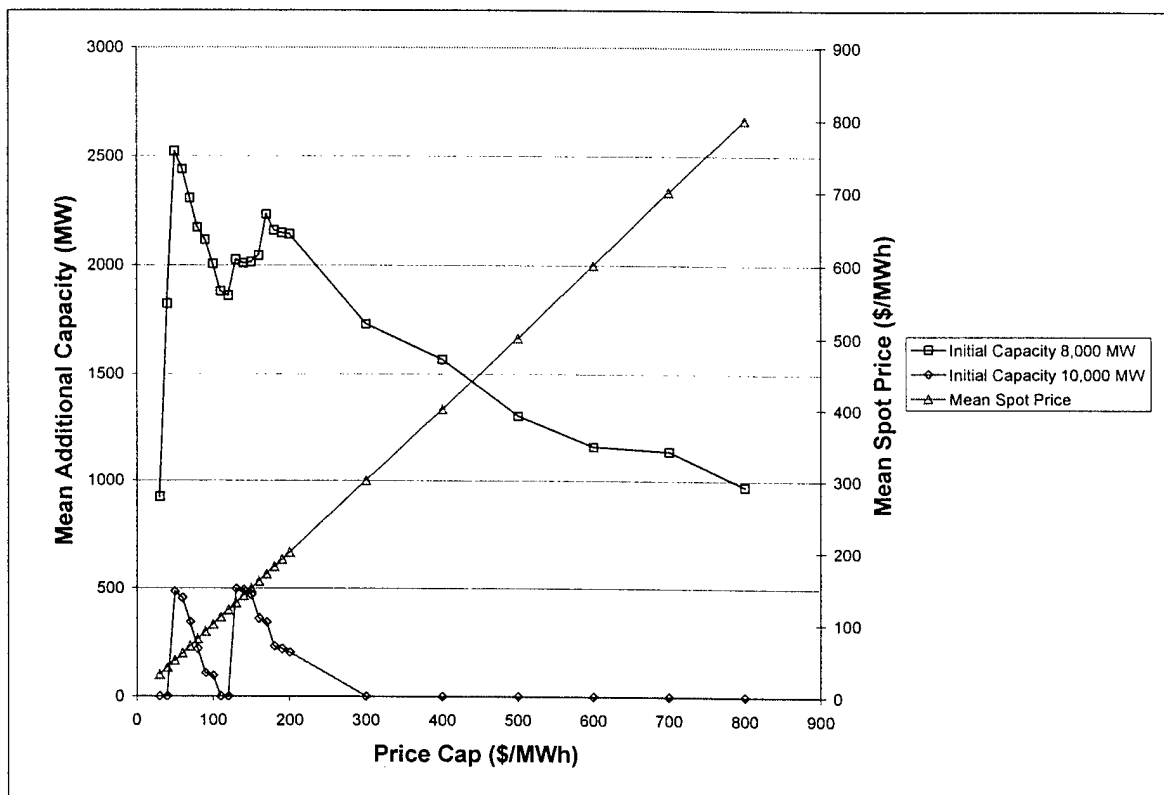


Figure 24. Effects of Price Cap on Monopolist

Cost Effect. At very low price cap levels, investment is inhibited when the price cap  $p_c$  is below average costs. In this case, no investment will be made because it is not profitable to invest in capacity that can never cover its average costs.

Demand Effect. This effect decreases investment as the price cap is increased and is caused by the fact that increased cap levels lead to lower levels of investment as a result of the demand-side response to higher prices. This effect is observed in each load

duration curve segment. Since demand is inelastic and the demand curve is iso-elastic, the monopolist facing a price cap will always choose to produce at:

$$L_{j,t}^* = (D_t + D_j^0) \cdot p_c^\varepsilon, \quad (3.2)$$

where  $L_{j,t}^*$  is the optimal production level for load duration curve segment  $j$  and time period  $t$ , given a price equal to the price cap  $P_c$ . The fact that the monopolist always produces at a quantity so that price is equal to the price cap  $P_c$  can be seen in Figure 24.

$D_t$  represents the demand shift parameter in time period  $t$ ,  $D_j^0$  represents the initial demand shift parameter level for load duration curve segment  $j$ , and  $\varepsilon$  is the price elasticity of demand. This production level is always optimal, because given that demand is inelastic, the marginal revenue of increasing output above  $L_{j,t}^*$  is zero. For output levels less than  $L_{j,t}^*$ , the monopolist will be forced to accept the price cap price for all quantity levels. Therefore, for these levels of output, it always benefits the monopolist to increase her output to  $L_{j,t}^*$ . Additionally, as the level of the price cap increases, this effect will reduce output level and result in a corresponding decrease in investment.

$$\frac{dL_{j,t}^*}{dp_c} = \varepsilon \cdot (D_t + D_j^0) \cdot p_c^{\varepsilon-1} < 0. \quad (3.3)$$

This effect exerts its influence across all price cap levels and is responsible for the decrease in capacity as the cap moves from \$60/MWh to \$120/MWh as well as the eventual decrease in capacity as the cap exceeds \$170/MWh. Total capacity will

approach zero as the price cap approaches infinity because the monopolist may set an unconstrained price on an infinitesimal quantity of energy.

Peak Load Effect. This effect opposes the demand effect and results in higher levels of investment as the price cap increases. This increase exists because, at lower price cap levels, it is not profitable for the monopolist to invest in enough capacity to meet  $L_{j,t}^*$  for peak and near-peak load duration curve segments because these levels of demand only occur a small percentage of the year. Therefore, total investment increases as the level of the price cap increases. This effect explains the second increase in capacity as the price cap increases from 120 to 170. The peak load effect is similar in direction to the cost effect; however, it is relevant for higher price cap levels than the cost effect. This is observed because prices greater than \$120/MWh are needed to justify capacity investments that can only be utilized for only a small portion of the year. A secondary result of this effect is that as the price cap level increases, the monopolist invests in a greater percentage of peaking technologies to meet peak loads.

The social welfare maximizer's response to price caps differs significantly from the monopolist's with respect to investment level and the overall effect on price. Mean additional capacity levels in year 10 along with the mean average yearly price for the social welfare maximizer are graphed in Figure 25. Mean additional capacity is listed on the left-hand axis and mean price is shown on the right-hand axis. Capacity levels increase monotonically as the cap level increases rather than bi-modally as in the monopolistic scenarios.

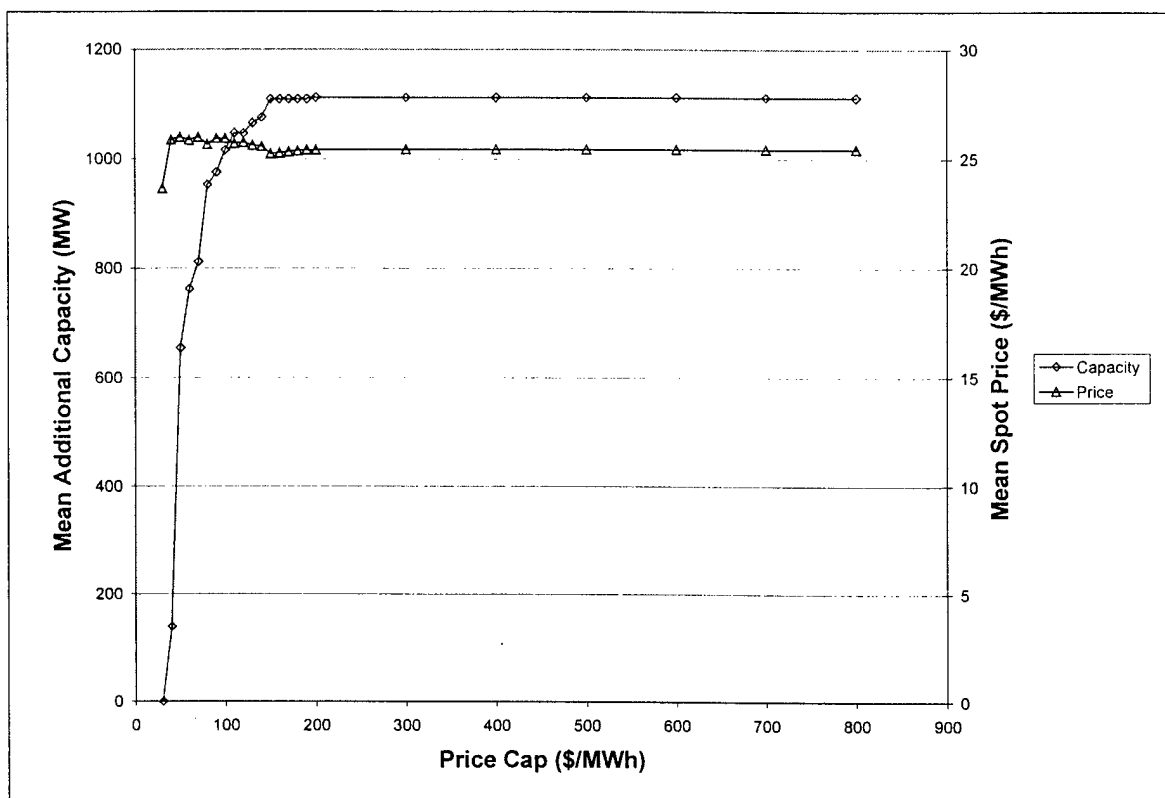


Figure 25. Effects of Price Cap on Social Welfare Maximizer

This results from elimination of the demand effect, which is not applicable to the social welfare maximizer because he will not restrict output for the purpose of increasing spot price. Both cost effects and peak load effects are active, thus contributing to the monotonic rise in investment with the increasing cap level. In this graph we also see that once the price cap is above the maximum average unconstrained price of approximately \$200/MWh, increasing the cap has a negligible effect on investment and average price.

Figure 25 also illustrates that average prices are slightly higher for the lower price cap levels compared with the higher price cap levels. This results from the dynamic

effects of price caps on investment. Since lower cap levels inhibit investment, when there is a lower price cap, spot prices hit the cap a greater percentage of the year than when there is a higher price cap.

This effect is illustrated in Figure 26 which compares prices across load duration curve segments for runs with no price cap, a price cap set to \$100/MWh, and a price cap set to \$50/MWh. The percentage of the year that each load duration curve segment is in effect is also plotted in Figure 26 to show the impact of each load duration curve segment price on the average yearly price.

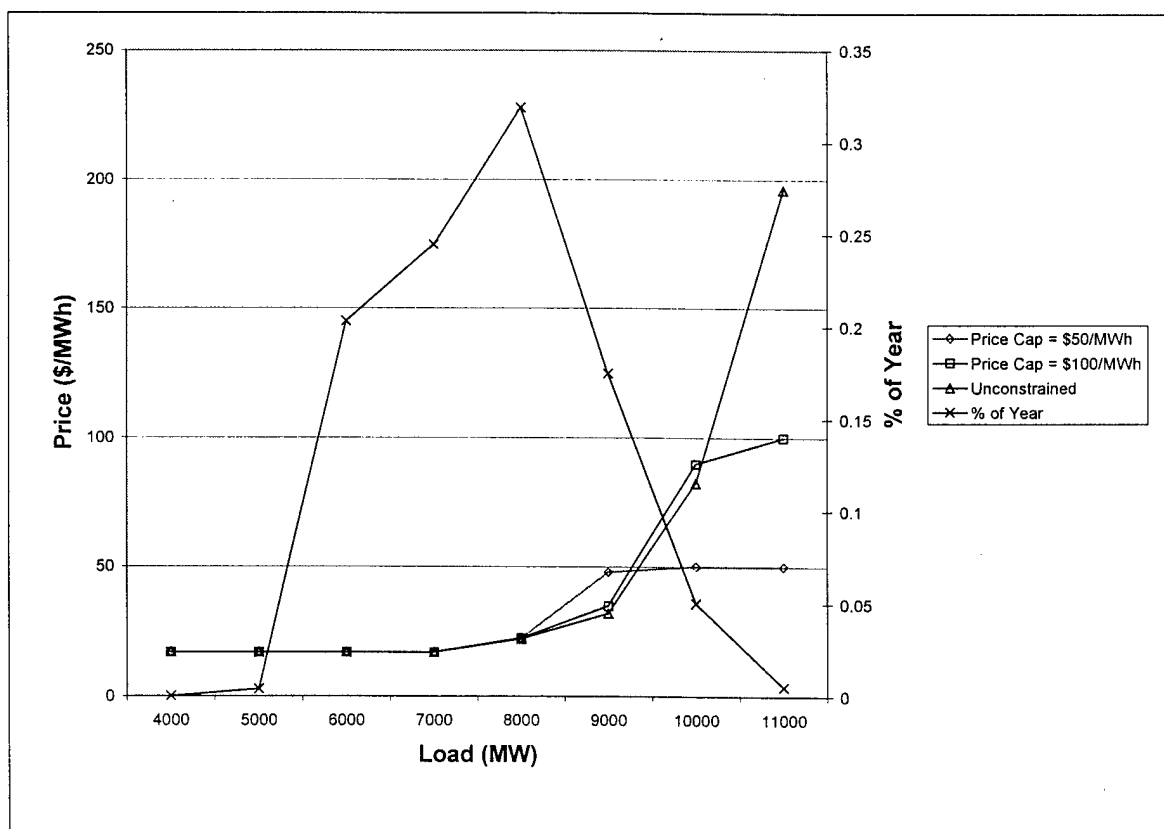


Figure 26. Effect of Price Caps on Price in Each Load Duration Curve Segment

Even though the peak price for the unconstrained scenario is significantly higher than the peak price in the price cap scenarios, the price at lower demand levels, most notably 9000 MW, is lower for the unconstrained scenario than for the price cap scenarios. Since this lower demand level occurs a greater percentage of the year, it has a relatively greater effect on average price.

The effects of the price cap on mean investment and mean price for both the social welfare maximizing and monopolistic perspectives are summarized in Table 6.

Table 6. Summary of Price Cap Effects

	Modeling Perspective	
	Social Welfare Maximizing	Monopolistic
Capacity	Monotonically Increasing with Cap Level	Bimodal
Price	No significant effects	Monotonically Increasing with Cap Level

The effects of price caps on investment level and spot price vary significantly based upon the modeling perspective.

### 3.6 Sensitivity of Peak Price to Demand Elasticity

This section of the essay examines the sensitivity of peak price to the elasticity of demand. This form of sensitivity analysis is conducted rather than running the capacity subsidy and price cap scenarios with different elasticities because those analyses may be difficult to interpret. When the absolute value of elasticity increases, lower levels of investment do not necessarily correspond to lower levels of generation adequacy because

when the elasticity increases, the underlying demand curve also changes. This analysis provides insight into the manner in which the previous results would be affected by varying demand elasticity.

The social welfare-maximizing model is run for elasticities ranging from -0.1 to -0.9 in increments of 0.1. For each elasticity value, demand curves are re-calculated using the “anchor point” technique discussed in Section 2.5.2. As in the previous sections, year 10 is selected for detailed analysis. Similar results exist for other years. Peak prices for each elasticity value are graphed in Figure 27.

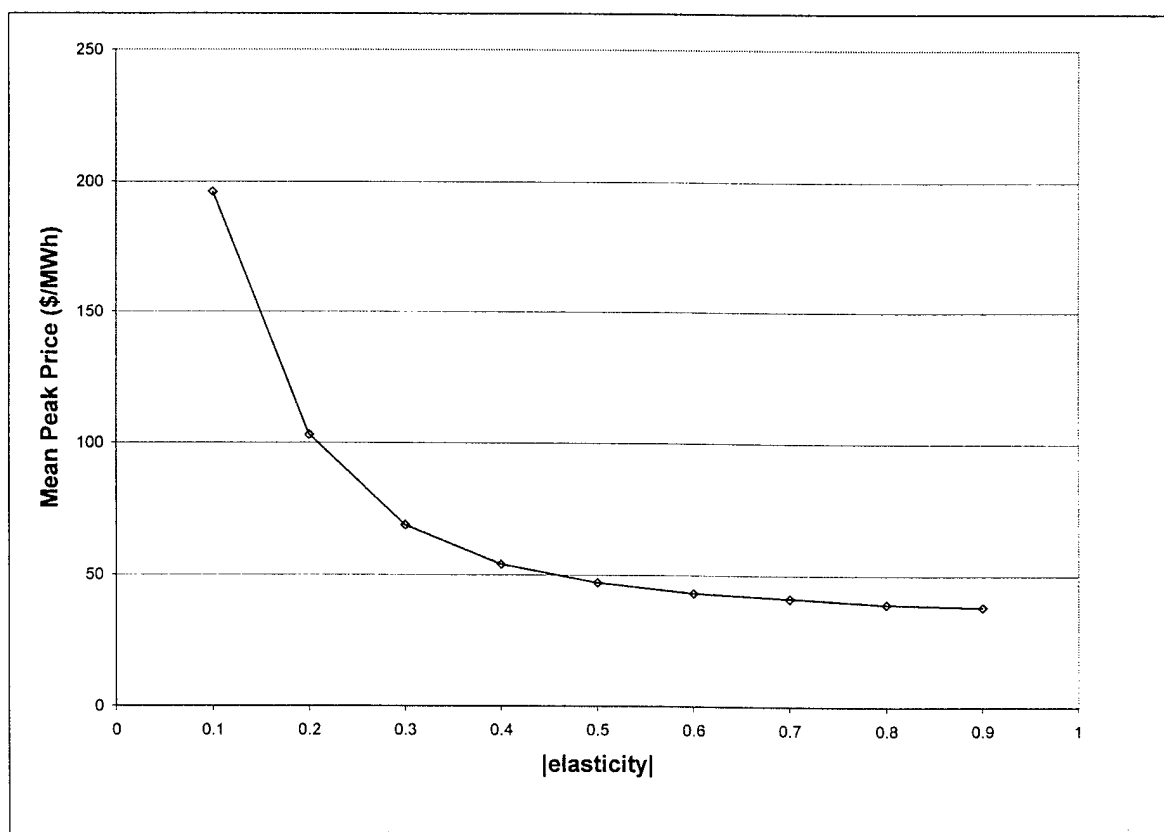


Figure 27. Effect of Elasticity on Peak Price

These prices decrease as elasticity increases and load is reduced based upon a demand-side response.

These results have implications for both the capacity subsidy and price cap results. Since peak loads can be curtailed by increasing the price elasticity of demand, the need for a capacity subsidy or price cap to curtail price volatility is lessened given more elastic demand. Similarly, the argument for capacity subsidies to ensure generation adequacy at peak loads is weakened if consumers can respond to price. Finally, since higher elasticities result in lower peak prices, the point at which price caps become nonbinding for the social welfare maximizer decreases as demand elasticity increases. This occurs because price caps above the peak price are nonbinding. These results also support the position of Wolak *et al.* (1999) that price caps are only needed during the transitional period between regulation and restructured competitive markets. Once mechanisms for demand-side responses exist, price caps can be removed.

### 3.7 Conclusions & Policy Implications

This essay demonstrates that the design of a restructured electricity market can significantly impact long-run investment behavior and electricity spot prices. This essay analyzes both capacity subsidies and price caps and determines their effect on investment level and spot market prices.

The results show that capacity subsidies act to reduce market volatility at the expense of increasing total electricity prices. However, as is discussed by Singh and

Jacobs (2000), capacity subsidies are probably not an efficient means of reducing volatility because they implicitly assume that all customers have similar risk preferences since market volatility is curtailed uniformly for all customers.

Therefore, as is discussed by Singh and Jacobs (2000) and Graves *et al* (1998) forward markets are preferable to capacity subsidies for the purpose of reducing price volatility. Forward markets provide customers who prefer not to risk price spikes the option to pay increased premiums to insure themselves against the possibility of these spikes. Similarly, those customers who are willing to accept price risk are rewarded through lower average prices. Another alternative that allows customers to manage risks associated with volatile prices is the use of derivatives such as options.

The model demonstrates that capacity subsidies increase overall levels of investment, which may positively impact reliability. However, this increase in reliability is applied uniformly to all customers despite evidence that not all customers desire the same level of reliability. For example, a hospital would most likely be willing to pay more for uninterrupted service than a residential electricity user. Additionally, hospitals and other large customers that require uninterrupted service may opt to achieve security through investments in distributed generation for use during peak periods. Therefore, an alternative approach that is also discussed by Graves *et al* (1998) involves letting customers self select their level of reliability by allowing some customers to sign up for interruptible service during peak loads in return for reduced rates. This policy would

create a demand-side response to price that would ensure market clearing for all levels of demand as well as reduce price volatility in the spot market.

One caveat to these policy suggestions is that the added reliability provided by a capacity subsidy may be worth its cost during the transitional period from regulation to restructuring. If an inelastic demand exceeds capacity, it may be impossible to determine an equilibrium spot market price if mechanisms for a demand side response to price are not yet in effect. Furthermore, this sort of situation can result in system failure due to the requirement that electricity systems must instantaneously balance supply and demand. One drawback to instituting transitional policies is that due to path dependence they may become locked in place. For example, LSEs may utilize political processes to keep capacity subsidies in place after they are needed.

Even though a direct equivalence does not exist between capacity subsidies and reserve requirements in the model presented in this essay, the investment response to reserve requirements likely would be similar to that seen from capacity subsidies. If so, reserve requirements for example as implemented in the PJM, would also increase investment and reduce price volatility at the expense of increasing the average total price of electricity.

Price caps have significantly different effects on investment behavior compared with capacity subsidies. These effects differ between monopolistic and social welfare maximizing scenarios. In the case of social welfare maximization, which approximates a competitive outcome, price caps do nothing to reduce average prices and may instead

increase average prices because of their deleterious effect on investment. In these situations price volatility is curtailed; however, the reduction in volatility comes at the expense of the social costs associated with the need to shed load during peak demand periods.

In the case of a monopoly supplier, price caps are required to limit market power. In the absence of price caps, the monopolist can increase prices without limit given the assumption of inelastic demand. While a price cap is necessary, it is difficult to determine the ideal cap level due to the bimodal response of investment to price. This bimodal outcome results from a combination of the cost and peak load effects that act to increase investment with higher price caps and the demand effect which inhibits investment for higher price cap levels.

The results from the extreme cases of monopoly and social welfare maximization can be integrated to develop policy insights for the State of Colorado, which is currently considering restructuring. Quick (2000) shows that the dominant firm in the Denver metropolitan region, Public Service Company of Colorado (PSCO), may have monopoly power for up to 54 percent of the year. Therefore if Colorado were to restructure, some form of price cap would be necessary to limit PSCO's mark-ups during these periods. However, it is important to ensure that this cap is set high enough so that no significant negative effect on long-run investment is realized. Results for the social welfare maximizer show that any price cap over \$200/MWh would have an insignificant negative impact on investment due to peak-load effects. Similarly, results from the monopolistic

perspective suggest that the cap should not be set significantly higher than \$300/MWh as a result of the demand effect. As market power is reduced and a demand-side response to price develops, any instituted price cap could be raised and ultimately phased out.

Future extensions to this research should explicitly consider cases of imperfect competition through the development of a multi-agent RL model. This would allow for a more accurate representation of the actual market structure in most locations.

Additionally, future research could experiment with finer discretizations of the load duration curve to more accurately determine the effects of capacity subsidies and price caps. Finer load duration curve segments would be especially helpful for accurately measuring the effect of higher cap levels on price. These discretizations would allow extremely high demand days that only occur once every several years to be incorporated into the model. On days such as these, price caps that were nonbinding for this model may in fact be binding. Finally, the model could be extended to account for the social costs associated with load shedding and to allow the model to estimate actual system reliability for each load duration curve segment. This would allow the model to estimate the welfare implications of proposed policies rather than simply the effect of proposed policies on investment and price.

## Chapter 4

### THE EFFECT OF UNCERTAIN TAX POLICY ON INVESTMENT IN WIND POWER

#### 4.1 Introduction

Many economists believe that wind power possesses the attributes of a positive externality (Cox, Blumstein, and Gilbert 1991; Mintzer, Miller, and Serchuk 1996). This belief is motivated by the fact that investment in renewables can offset investment in traditional fossil fuel-based generation and thereby reduce the pollution related social costs associated with fossil fuel generation (DOE 1997; Gipe 1995, 423). Several studies have attempted to quantify the social costs associated with electric power fossil fuel emissions (Desvousges, Johnson, and Banzhaf 1994; Rowe, Bernow, and White 1995; Freeman and Rowe 1995). In addition, some have argued that wind power provides “energy security” by reducing reliance on imported oil for power production and diversifying the generating fuel base (Cox, Blumstein, and Gilbert 1991, 348).

Since the market will underprovide goods that possess the characteristics of a positive externality, a wind power subsidy or fossil fuel emissions tax may be justified (Nijkamp 1977, 45). In addition to the externality justification, others have pushed for government support of wind power based upon the argument that wind power is an infant industry which needs to be fostered until it can stand on its own (Cox, Blumstein, and Gilbert 1991, 366). Therefore, numerous wind power subsidy programs and emissions

taxes have been enacted in the United States and other countries in order to encourage the development and use of wind power and other renewable technologies. A great deal of research has focused on the effects of both subsidies and taxes that are intended to internalize the externalities associated with electrical power generation (Bernow, Biewald, and Marron 1991; Burtraw, Palmer, and Krupnick 1993; Palmer and Dowlatabadi 1993).

Rather than attempting to determine what level of tax/subsidy is efficient or analyzing the merits of a given policy, this essay focuses on the effect of policy uncertainty on investment in wind power. Specifically, uncertainty over the enactment or repeal of investment tax credits (ITCs) and production tax credits (PTCs) is investigated. The effect of policy uncertainty on wind power investment is relevant because public policy toward wind power has historically been highly variable and prospective wind power investors face considerable uncertainty relating what policies will be in effect in the future. This research extends the literature relating to the effects of uncertain tax policy on investment behavior by focusing on uncertain tax policies that apply to only one technology from a group of substitutable technologies (Dixit and Pindyk 1994; Hassett and Metcalf 1999).

The remainder of the essay is organized as follows: Section 4.2 discusses the history of public policy relating to wind power in the United States as well as several proposed policies relating to wind power. Section 4.3 summarizes the relevant literature on the effects of policy uncertainty on investment behavior. Section 4.4 presents a

reinforcement learning-based model of generation investment under demand and tax policy uncertainty. This model is used to analyze how anticipation of the enactment or repeal of an investment tax credit or production tax credit will affect investment in wind power. Section 4.5 summarizes the results from this model and Section 4.6 provides concluding remarks as well as a discussion of policy implications from this work.

#### 4.2 Public Policy History Pertaining to Wind Power

Prior to the 1970s, no significant federal or state policies were implemented to increase the rate at which wind power was adopted by the United States electric industry. However, concerns over reliance on imported oil, sparked by the Arab Oil Embargo of 1973, increased the importance placed upon “energy security.” This term refers to the public good characteristics of using a diverse set of fuels to hedge against the macroeconomic impacts associated with price shocks in one type of fuel. These price shocks are more likely for fuels that are imported from unstable regions such as the Persian Gulf (Fox-Penner 1997, 357).

These concerns, along with more traditional environmental considerations, motivated the passage of the National Energy Act (NEA) in 1978 (Cox, Blumstein, and Gilbert 1991). This legislation called for a 15 percent ITC on all wind power investments, which supplemented an existing 10 percent federal ITC that applied to all classes of investments. Another provision in the NEA legislation called for \$100 million dollars in cooperative agreements, grants, and subsidized loans to further spur

development of the United States wind power industry (Cox, Blumstein, and Gilbert 1991, 354). The 15 percent ITC from the NEA was phased out in 1985.

Another related piece of legislation enacted in 1978, as a portion of the NEA, was the Public Utilities Regulatory Policy Act (PURPA). This legislation, which represented a first step toward wholesale electricity competition, created a mechanism for owning and operating power plants in which the owner was exempt from price regulation. Owners designated as qualifying facilities (QFs) could sell electricity to regulated power companies whom would then sell power to their customers. PURPA not only allowed for the sale of power to regulated firms, but also required these regulated utilities to buy power from QFs in their region based upon "avoided costs" to the utility. To qualify as a QF, plants needed to utilize either cogeneration or provide power via renewable sources such as wind power (Fox-Penner 1997, 15).

In addition to this federal activity, California adopted a state ITC in 1978 which was applied on top of all federal tax credits resulting in an aggregate 50 percent tax credit for wind investors in California. This ITC was eventually phased out in 1987. Additionally, California's implementation of PURPA required that "avoided cost" calculations, which set prices between QFs and regulated utilities, were made by the California Public Utilities Commission (CPUC). The CPUC often set these prices under terms favoring the QF's (Cox, Blumstein, and Gilbert 1991, 355). The combination of the NEA, California's Wind Power ITC, and California's method for implementing PURPA caused a boom in wind power investments in California by independent

investors (Righter 1996, 209). From 1982 through 1985, wind power capacity in California grew from 7 MW to 1,141 MW. The growth in wind power capacity was so significant that by 1987, California produced 87 percent of the total world wind power (Cox, Blumstein, and Gilbert 1991, 356).

However, by the early 1990s the California-led revolution in wind power investment had subsided as a result of the removal of federal and state tax incentives as well as falling natural gas prices resulting from deregulation of the natural gas industry. The fall in natural gas prices impacted the “avoided cost” calculations which directly affected the profitability of the QFs. In 1983, the CPUC set avoided costs in California at approximately 8 cents per kWh. In contrast, avoided costs during the mid-1990s were roughly 5 cents per kWh—a figure that motivated several turbine shutdowns (Righter 1996, 222). The CPUC recalculated avoided costs during the mid nineties because QF contracts called for 10 years of fixed prices followed by floating rates for the next 20 years. As a result of these avoided cost recalculations and numerous technical failures, investors deactivated over 230 MW of California’s installed wind capacity (Gipe 1995, 475; Stuebi 1999).

Finally, negative publicity relating to wind power reduced public support for the technology. There were numerous reports that wind power was responsible for bird deaths—specifically, the death of raptors in the Altamont Pass area of California. These claims were somewhat exaggerated and there is little evidence that the bird death rate from collisions with wind turbines is significantly greater than the bird death rate

produced by any large structure (Benner 1992; Orloff and Flannery 1992). In addition, investors installed many of the least reliable turbines along Interstate 580 over Altamont Pass. Since these turbines were often idle, especially during the peak traffic periods, the public perception that wind power was an ineffective technology was reinforced (Gipe 1995, 275).

More recently, the United States has experienced a nation-wide rebirth in wind power investment brought about by technological improvements in wind turbine technology. These technological improvements have reduced costs and improved the reliability of wind turbines. From 1980 to 1999 costs fell from over 25 cents per kWh to approximately 4 cents per kWh (Gipe 1995, 233; Steve 1999). A federal Production Tax Credit (PTC) of 1.5 cents per kWh from December 31, 1993 through June 30, 1999 has also aided this resurgence. This PTC was recently extended through December 31, 2001. The PTC calls for a tax credit on all generated wind power that originates from turbines that were commissioned while the legislation was in place. The credit remains in effect for the first 10 years that a wind turbine is operating and is only valid if the wind turbine is located within the United States and electricity is sold to an unrelated party. The primary motivation for this legislation is to keep wind power competitive as more states convert from regulated monopolies to restructured electricity markets (Steve 1999).

Figure 28 shows total the total installed wind capacity in the United States from 1981 through 2000 (AWEA 2000c). The first significant increase in wind power investment occurs during the "California Wind Boom" from 1982 through 1987. The

second major increase occurs from 1997 through 1999 due to recent technological improvements as well as the PTC.

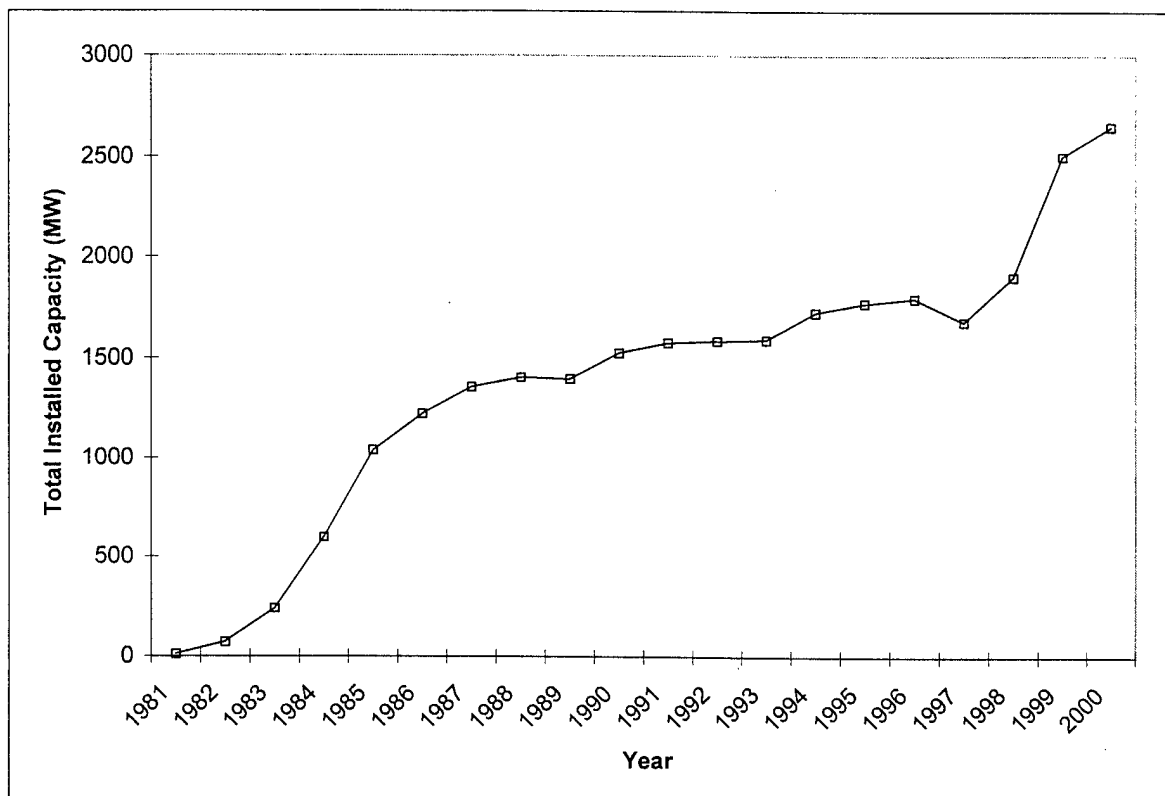


Figure 28. Total Installed United States Wind Power Capacity

In addition to the previously described tax policies, several other types of policy initiatives including a renewable portfolio standard (RPS), system-benefit charges (SBC), and green pricing have been enacted at the state level and are currently under consideration at the federal level. A RPS requires that a fixed percentage of all generated electricity originate from renewable technologies (Awerbuch 2000). This sort of program may have significant advantages over technology-specific subsidies, such as the PTC or

ITC on wind power, because it allows for the market to determine an efficient mix of renewable technologies. This mix is determined based upon the technical merits and cost-effectiveness of each technology. Another benefit to a RPS is that it allows each region to invest in the technologies that are most appropriate for that region (AWEA 2000a). Several states including Connecticut, Maine, Nevada, New Jersey, Pennsylvania, Texas and Wisconsin have already adopted a RPS while several federal restructuring bills contain provisions for a national RPS (AWEA 2000a). Proposed federal legislation ranges from a 3 percent to a 10 percent renewable production requirement by the year 2010. The most stringent proposal, in Senate Bill 1369, calls for a yearly  $\frac{1}{2}$  percent increase in the RPS until 2005 followed by a 1 percent yearly increase after 2005 through the year 2020. This would result in a 20 percent standard by the year 2020 (AWEA 2000a).

System benefit charges (SBCs) impose a per-MWh fee on all demanders. These fees are collected and then distributed to owners of renewable generation. These SBC policies have been implemented in California, Connecticut, Illinois, Massachusetts, Montana, New Jersey, New Mexico, New York, Pennsylvania, and Rhode Island. Advocates for SBCs suggest that they give policy makers more flexibility in allocating funds to support infant industries. For example, 10 percent of California's SBC funds are dedicated to "higher cost emerging technologies" such as photovoltaics (Wiser, Porter, and Clemmer 1999). Some economists do not believe that infant industries should be subsidized. They argue that as long as capital markets are efficient, then investors will

finance industries with the prospect of high returns in the future. An example of this behavior was observed in the biotechnology industry, which attracted hundreds of millions of dollars of capital years before any profits were realized (Krugman and Obstfeld 1992, 255). Another argument against government funding for infant industries is that it is unlikely that the government possess enough information to pick “winning” industries.

Green pricing allows residential and industrial customers a pay premium for their electrical power in order to support renewable electric generation. Green pricing is currently being used in Arizona, California, Colorado, Florida, Hawaii, Michigan, Minnesota, Nevada, Oregon, Texas and Wisconsin. In California, green power from in-state generation is subsidized so customers can buy it at a discount. However, there is considerable uncertainty as to whether this particular renewable customer credit will be extended beyond the year 2002. Also, it is possible that the level of the subsidy may be reduced prior to 2002 (Byrne 1999).

Another proposal, sponsored by the American Wind Energy Association, calls for a 30 percent federal tax credit for individuals and businesses that employ wind turbines with a total rated capacity of less than 50 kW. Agricultural and residential users who are geographically removed from the power grid primarily use this class of wind turbine. While these users are small in number, this proposal has the potential to impact overall pollution levels because potential users of small wind turbines currently employ highly polluting diesel generation (AWEA 2000b).

One common theme throughout these policy changes and proposed policy changes is that the type and duration of public policy toward wind power has changed over time. It is likely that public policy toward wind power will continue to change in the future based upon the presidential administration, the composition of federal and state legislatures, and volatile fossil fuel prices. Additionally, uncertainty relating to enactment and enforcement of environmental agreements such as the Kyoto protocol will indirectly affect public attitudes toward wind power investments. These factors create a significant source of uncertainty for investors considering wind power investments. Section 4.3 summarizes research related to the effect of policy uncertainty on investment behavior.

#### 4.3 Literature on Investment Under Policy Uncertainty

A large percentage of theoretical research on policy uncertainty has focused on the area of ITC uncertainty. Dixit and Pindyk (1994) employ a firm level model to demonstrate two basic effects of tax policy uncertainty. When an ITC is in place and there is the probability that the credit will be removed, firms will accelerate their investment decisions in order to take advantage of the credit before it is removed. Conversely, Dixit and Pindyk show that when an ITC is not in place, but the probability exists that an ITC will be enacted in the future, the level of investment is decreased due to the increased option value of waiting for potential ITC enactment.

Hasset and Metcalf (1999) expand these results to determine implications for total capital stocks based upon differing forms of policy uncertainty. They conclude that the structure of the stochastic process describing tax policy uncertainty determines the aggregate effects of uncertainty on the total capital stock. If a nonstationary process such as geometric Brownian motion is assumed, they demonstrate that aggregate effects will be negative. Conversely, a stationary process, such as a Poisson jump process, will increase aggregate capital levels. However, the strength of this investment increasing property of stationary tax policy uncertainty is reduced if policy parameter movements are negatively correlated with price movements.

Another key result of Hasset and Metcalf (1999) is that increased tax policy uncertainty—regardless of its form—will result in decreased tax revenue because uncertain tax policy acts as an “implicit subsidy” to firms. This subsidy originates from the intertemporal substitution of investment from time periods with lower subsidy levels to time periods with higher subsidy levels. Therefore, they argue that it is in the best interest of the government to pursue tax policy stability. Also, Bizer and Judd (1989) employ a general equilibrium framework to show that tax policy uncertainty will always lead to a reduction in social welfare.

Other related empirical research focusing on macroeconomic policy uncertainty in developing countries shows that policy uncertainty is negatively correlated with aggregate levels of investment. However, this negative effect is somewhat mitigated by

policy persistence—the length of time that policies remain in effect (Aizenman and Marion 1993).

Related research pertaining specifically to environmental policy uncertainty has looked at the effect of an uncertain transferable discharge permit policy on investment in wastewater treatment plants. Results show that if there is doubt over whether or not future discharge permit trades will be permitted, the number of trades that actually will be made falls. This effect results in a reduction in the benefits realized from transferable permit programs. The overall implications of these results are that discharge-trading programs may not achieve their intended objectives in environments characterized by high levels of policy uncertainty (Leston 1992).

In the area of generation investment decisions, Teisberg-Olmsted (1993) determines that regulated utilities facing uncertainty over future allowable rates of return will favor smaller, shorter lead-time investments. This result stems from the added flexibility that this class of investment provides, given that future regulatory conditions are unknown at the time of the investment decision.

The research in this essay differs from the majority of the previously described work on tax policy uncertainty in that its focus is on tax policy uncertainty applied to one specific technology rather than generic investment. This is significant because firms may substitute between wind power and classical investments, which may in turn exacerbate the effects of uncertainty. Policy uncertainty effects may be stronger when substitution opportunities exist between subsidized and unsubsidized technologies because firms

anticipating an ITC enactment may defer all wind power investment until an ITC is enacted. A compensating or partially compensating increase in classical investments may be used to offset this decrease in wind power investment. Similarly, firms could reduce investment in classical technologies to compensate for an increase in wind power investment if an ITC removal is anticipated. A second unique characteristic of this essay is the use of reinforcement learning (RL) to model the effects of policy uncertainty. Using RL facilitates the multiple technology model presented in this essay since multidimensional state transitions would be difficult to define explicitly using classical techniques.

#### 4.4 Model

Basic modeling assumptions are presented in Section 3.4.1, the mathematical structure of the model is presented in Section 3.4.2, cost and technical data are presented in Section 3.4.3, and Section 3.4.4 describes the various scenarios that are considered.

##### 4.4.1 Assumptions

All assumptions from the basic model presented in Section 2.4.1 of the first essay apply to this model. Additionally, the following assumptions are added:

Technologies. The agent may determine its investment portfolio from an action space comprised of two technologies. These technologies are (1) a composite technology comprised of a 50-50 mix of combined cycle (CC) and combustion turbine (CT)

generation and (2) wind power. By assumption, it is impossible for the firm to invest in combinations of CC-to-CT other than 1-to-1. Therefore, the agent's CC capacity will always equal its CT capacity.

Social Welfare Maximization by an "Independent" Agent. This model assumes that an independent SW maximizing agent makes a long run investment decision every year concerning the level of investment in each technology. The agent also makes short-run dispatch decisions for each segment of the load duration curve in order to maximize social welfare. For the purposes of this model, social costs (e.g., fossil fuel emissions) resulting from dispatch of the classical technology portfolio are not considered. The model ignores the fact that the ITC and PTC may be financed through higher taxes that would reduce consumer surplus. Finally, the social welfare maximizing agent has no knowledge of future realizations of the stochastic component of demand uncertainty or tax policy uncertainty.

Energy is a Homogeneous Good. Consumers treat power from classical and wind generation sources equally and there is no "green premium" or greater willingness to pay for wind power.

Independence Between Demand and Policy Uncertainty. It is assumed that the load duration curve grows based upon a discrete state random walk with drift. Also, the model assumes that policy uncertainty can be modeled using a Markov Chain. Therefore, every period there is a discrete probability that the policy in question will be implemented if it is not currently implemented and there is a discrete probability that the policy in

question will be removed if it is currently in place. In addition, the model assumes statistical independence between demand and policy uncertainty.

#### 4.4.2 Model Structure

The modeling framework utilized in this essay is similar in structure to the general modeling framework presented in Section 2.4 of the first essay. In this framework, a policy mapping from states to actions is determined such that expected discounted social welfare is maximized. Next, this optimal policy is utilized in the simulation module to determine mean levels of capacity across time. Only differences between the modeling framework presented in the first essay and the modeling framework used in this essay are discussed below.

The state space is 4-dimensional, defined by the capacity of the classical portfolio, the capacity of wind power, the demand shift parameter, and the policy parameter. The state space ranges from 10,300 MW to 10,900 MW of classical technology in 150 MW increments and from 100 MW to 480 MW of wind power in 20 MW increments. Wind power capacity is discretized in 20 MW blocks rather than 150 MW blocks based upon the size of the majority of wind farms in the Rocky Mountain Power Area (RMPA) (AWEA 1999). Demand shift parameter values range from 0 MW to 750 MW in 150 MW increments. An initial level of 100 MW is used for wind power to approximate actual wind power capacity in the RMPA (AWEA 1999). The initial classical capacity of 10,300 MW is assumed to ensure that the mean price in the first year of the simulation is

approximately equal to \$30/MWh, an approximation of the 1998 average wholesale electricity price (Stone and Webster 1998).

Finally, the policy parameter is either equal to 1 indicating that the policy in question is in place or 0 signifying that the policy in question is not in place. ITC and PTC policies are considered separately to avoid adding another dimension to the state space.

The action space allows for investment in 20 MW blocks of wind power and 150 MW blocks of classical generation. The maximum per-period wind investment is 40 MW and the maximum per-period investment in the classical portfolio is 300 MW. The maximum wind investment rate of 40 MW is justified based upon practical and computational considerations. First, engineering constraints limit the amount of total system capacity that can be comprised of wind power to roughly 5 percent (Putnam 1996). Additionally, the recent worldwide increase in demand for wind turbines has created a significant production backlog on wind turbines (Poulsen 1999). Computational considerations also contribute to the decision to bound maximum allowable wind investment at 40 MW. Reinforcement learning requires a rather parsimonious action space in order to keep run-times reasonable. Therefore, the only means to increase the maximum allowable level of investment, without increasing the size of the action space, involves increasing the block size on wind power investment. This approach is undesirable since only a small percentage of wind farms in the United States are greater than 20 MW (AWEA 1999).

Table 7 summarizes the 9 actions in the action space. The action space is larger than that used in the first essay because wind investments are small enough in total dollars to be made independently from the classical investment decision. Even though wind power investment in any given year will not likely offset a significant quantity of classical generation investment, it is likely that wind power investment across several years will be able to partially offset investment in the classical technology.

Table 7. Action Space

Action Index	Investment in CC/CT Mix (MW)	Investment in Wind Power (MW)
0	0	0
1	0	20
2	0	40
3	150	0
4	150	20
5	150	40
6	300	0
7	300	20
8	300	40

Equations of motion for capacity and the demand shift parameter are identical to those presented in section 2.4.2.4 of the first essay. The policy parameter  $P_t$  transitions based upon the following Markov Chain:

$$p(P_t = 1 | P_{t-1} = 0) = \lambda_0, \quad (4.1)$$

$$p(P_t = 0 | P_{t-1} = 0) = 1 - \lambda_0, \quad (4.2)$$

$$p(P_t = 0 | P_{t-1} = 1) = \lambda_1, \quad (4.3)$$

$$p(P_t = 1 | P_{t-1} = 1) = 1 - \lambda_1, \quad (4.4)$$

where,  $\lambda_0$  is the probability of transitioning from a state without the policy in effect to a state with the policy in effect and  $\lambda_1$  defines the probability of transitioning from a state with the policy in effect to a state without the policy in effect.

Finally, the model's reward structure is identical to that presented in section 2.4.2.5 of the first essay with the only difference being the manner in which the various tax policies are represented. The ITC reduces total investment costs per MW by a fixed percentage. In the case of the ITC, the benefit is received without regard to the way in which the technology is used. In contrast, the PTC rewards firms by subsidizing them for each MW of wind power that they dispatch. Thus, benefits from the PTC only accrue from the use of wind power, rather than from the act of investing in wind power.

This manner of implementing the PTC differs from the federal PTC that was enacted in 1993 because the actual PTC only provides a tax credit to firms that invest while the policy is active. The PTC is modeled differently from the enacted PTC for two reasons. First, computational requirements prohibit modeling of a policy similar to the actual PTC because this sort of policy requires an additional dimension in the state space to account for wind power investments made under the policy vs. those made when the policy is not in effect. A second rationale for modeling the PTC in this manner is to capture the investment effects of a policy that is applied without regard to the year in which an investment is made. While the actual PTC does not share this property, the proposed RPS does operate in this manner. The RPS involves a production requirement that is not related to the year in which renewable investments are made.

#### 4.4.3 Data

Cost and availability data for the classical technology portfolio and wind power are listed in Table 8.

Table 8. Cost and Availability Data

	Classical	Wind
Variable Cost ( $vc$ )	21.5 \$/MWh	1 \$/MWh
Fixed Cost ( $fc$ )	5,630 \$/MW	7,550 \$/MW
Investment Cost ( $IC$ )	478,500 \$/MW	1,000,000 \$/MW
Availability/Capacity Factor	0.9	0.5

Costs for the classical portfolio are the average of the costs for the CC and CT technologies since a 1-to-1 mix is of CC-to-CT is assumed. Cost values for wind data are based on discussions with representatives from New Century Energies who operate the Foote Creek and Ponnequin wind farms in Wyoming and Northern Colorado (Sulkko 1999). The capacity factor of 0.5 is used based upon a range of capacity factors from 0.2 to 0.6 that are reported in the literature (Cavallo 1995; DOE 1997). The capacity factor is equal to the percentage of the wind power's total rated capacity that will be available for dispatch over the course of a year. This factor is implemented in the model in a manner similar to availability and determines the total amount of capacity that is available for dispatch in each load duration curve segment. The term capacity factor is used in place of availability for wind resources because wind patterns, rather than maintenance requirements, are the largest determinant of the percentage of a turbine's rated capacity that is available for dispatch across a given load duration curve segment.

The capacity factor estimate of 0.5 was selected from the upper range of reported values to ensure that some investment in wind power technology would occur in scenarios with no tax subsidies in place. This assumption was made so that this study could evaluate both investment increasing and investment decreasing effects that originate from the expectation of policy addition or removal. If the agent did not invest in wind power without a tax subsidy in place, it would have been impossible to show the investment decreasing effect that originates from the expectation of an ITC. Higher or lower capacity factors should increase or decrease the total level of wind power investment; however, they should not significantly change the nature of the response to policy uncertainty. Also, it is important to note that capacity factors vary from site to site based upon wind conditions as well as turbine technologies. Therefore, assumption of a constant capacity factor across all new wind capacity additions is also a slight departure from reality.

#### 4.4.4 Scenarios

Prior to determining the effect of policy uncertainty on wind power investment, the individual effects of the ITC and PTC on investment are established. Scenario 1 compares investment in wind power with no ITC to investment with ITC levels ranging from 10 percent to 30 percent. This range is utilized since it bounds the 15 percent ITC that was implemented as a part of the NEA. Similarly, Scenario 2 compares the base

case with no policy to PTC levels of \$7.5, \$15 and \$22.5 per MWh. These values are utilized to bound the current PTC of \$15 per MWh.

Next, Scenario 3A models investment with no ITC policy in place and no expectation of a transition to an ITC policy ( $\lambda_0 = 0, \lambda_1 = 0$ ) and Scenario 3B models investment with no ITC policy in place when there is the expectation of an irreversible transition to a state with the ITC in place ( $\lambda_0 = .5, \lambda_1 = 0$ ). Since there is strong expectation of an ITC in the subsequent period, we would expect that Scenario 3B would deter investment compared with Scenario 3A. In the simulations for Scenarios 3A and 3B the initial state has no ITC in place and the policy is never enacted in order to illustrate the degree to which the expectation of an ITC can deter investment. Scenario 3C models a situation where an ITC is in place and there is no chance of removal ( $\lambda_0 = 0, \lambda_1 = 0$ ), and Scenario 3D investigates a situation where there is an expectation of an irreversible removal of the ITC ( $\lambda_0 = 0, \lambda_1 = .5$ ). Scenario 3D may encourage higher levels of investment compared with Scenario 3C because firms expecting an irreversible ITC removal should invest at a higher level to take advantage of the ITC before it is removed. In the simulation module for Scenarios 3C and 3D, the initial state utilizes the ITC and all of the subsequent states also have the ITC in effect to illustrate how the expectation of an ITC removal can increase total investment. For both the RL and simulations modules, Scenarios 4A through 4D are identical to 3A through 3D except they examine a PTC rather than an ITC.

Finally, Scenarios 5 and 6 determine the impacts of the respective ITC and PTC policies while varying  $\lambda_0$  and  $\lambda_1$  from 0.0 to 0.5 in increments of 0.1. Therefore, Scenarios 5 and 6 are each comprised of a total of 36 model runs. Scenario 5 uses an ITC of 10 percent and Scenario 6 utilizes a PTC of \$15/MWh. Investment behavior under the varying levels of policy uncertainty is measured by averaging wind investment actions from the optimal RL policy for states with the tax policy in effect and for states without the tax policy in effect. Therefore, the simulation module of the general framework is not utilized in Scenarios 5 and 6. This metric is used in place of the simulation module to interpret the results of Scenarios 5 and 6 so that the results are not sensitive to the policy in the initial simulation state.

Table 9 summarizes the conditions associated with each of the Scenarios 1 - 6:

Table 9. Summary of Scenario Conditions

Scenario	RL Module		Simulation Module				
	$\lambda_0$	$\lambda_1$	$\lambda_0$	$\lambda_1$	Initial Policy	Comparison Metric	Policy
1	0	0	0	0	ITC	Sim.	0%-30 % ITC
2	0	0	0	0	PTC	Sim.	\$0 - \$22.5 PTC
3A	0	0	0	0	No ITC	Sim.	10% ITC
3B	.5	0	0	0	No ITC	Sim.	10% ITC
3C	0	0	0	0	ITC	Sim.	10% ITC
3D	0	.5	0	0	ITC	Sim.	10% ITC
4A	0	0	0	0	No PTC	Sim.	\$15 PTC
4B	.5	0	0	0	No PTC	Sim.	\$15 PTC
4C	0	0	0	0	PTC	Sim.	\$15 PTC
4D	0	.5	0	0	PTC	Sim.	\$15 PTC
5	0 - .5	0 - .5	N/A	N/A	N/A	Ave. Pol.	10% ITC
6	0 - .5	0 - .5	N/A	N/A	N/A	Ave. Pol.	\$15 PTC

#### 4.5 Results

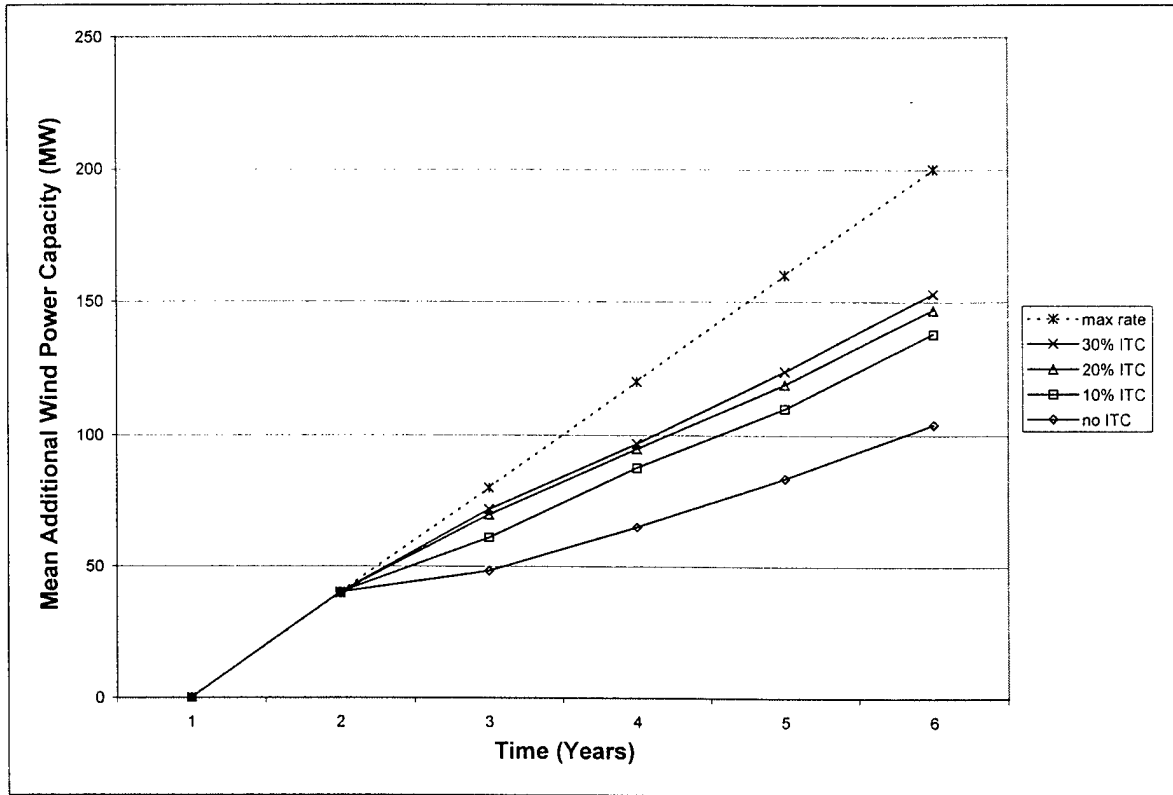


Figure 29. Varying ITC Level (Scenario 1)

Figure 29 shows results from Scenario 1 that demonstrate the mean levels of wind power investment for ITC levels ranging from 0 to 30 percent. As expected, the mean investment level increases with higher levels of the ITC, however, the marginal impact of the ITC decreases as the ITC level increases. Figure 29 also plots the maximum rate at which the model could invest in wind power for comparative purposes. This maximum investment rate originates from the upper bound on wind power investment imposed on the action space. Figure 29 through Figure 32 only graph new wind power capacity

additions and do not include the original 100 MW of wind power included in the state space.

Figure 30 shows mean aggregate wind power capacity investments from the simulation module for Scenario 2. These results also show increased investment at higher PTC levels with a decreasing marginal effect of the PTC. The maximum allowable investment rate is also plotted for comparative purposes. It is likely that investment levels for the higher PTC levels would exceed these levels if not for the upper bound of 40 MW imposed on the model's action space.

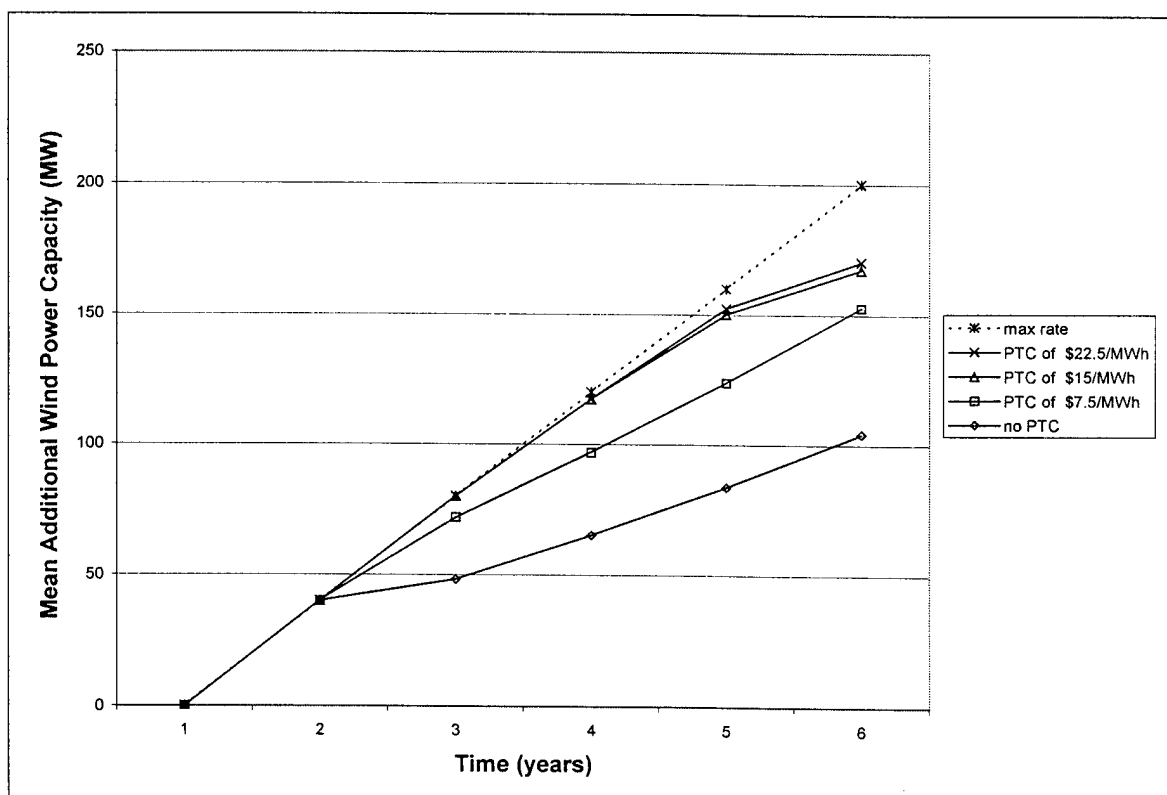


Figure 30. Effect of Varying PTC (Scenario 2)

Figure 31 demonstrates the results from Scenarios 3A through 3D. In Scenario 3B, the expectation of an ITC enactment significantly reduces investment in each period because the agent is waiting for the ITC to be enacted prior to investing. Similarly, in Scenario 3D we see an increase in investment when the ITC is in place and there is an expectation of a transition to a state with no ITC in effect. Note that the investment postponing effect from the expectation of ITC removal is stronger than the investment accelerating effect resulting from the expectation of ITC enactment. The differing strengths of these effects are consistent with the results of Dixit and Pindyk (1994).

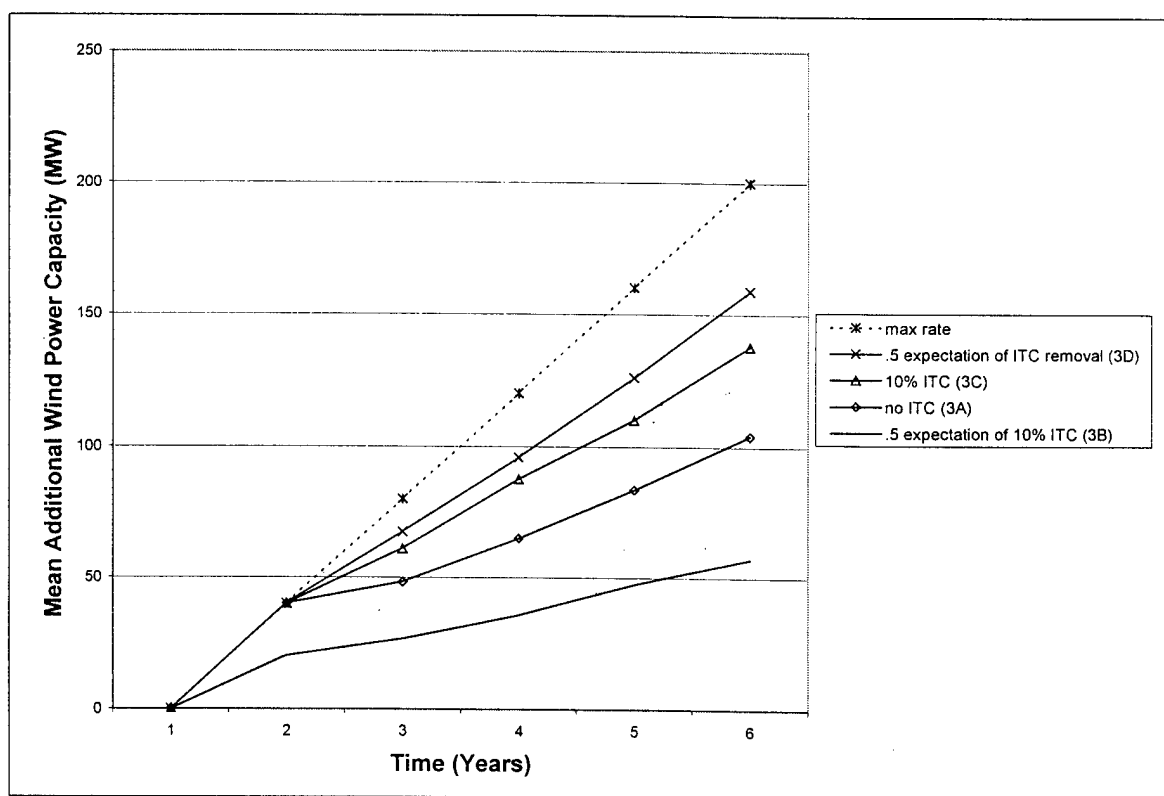


Figure 31. Effect of Expected ITC Removal or Addition (Scenario 3A – 3D)

Results from Scenarios 4A through 4D are shown in Figure 32. These results show that investment increases when the agent expects to transition permanently to a state with the PTC (4B) and the investment level decreases when permanent removal of the PTC is expected (4D). Both of these effects are opposite in direction to the effects seen with the ITC. This opposing result can be explained by the nature of the incentive. Unlike the ITC, that only benefits firms at the time of the investment decision, the PTC benefits firms in all periods after the investment is made provided that the capacity is used to generate electricity. Therefore, a firm's expectation that the PTC will be in effect in future periods will influence whether or not they invest in wind power.

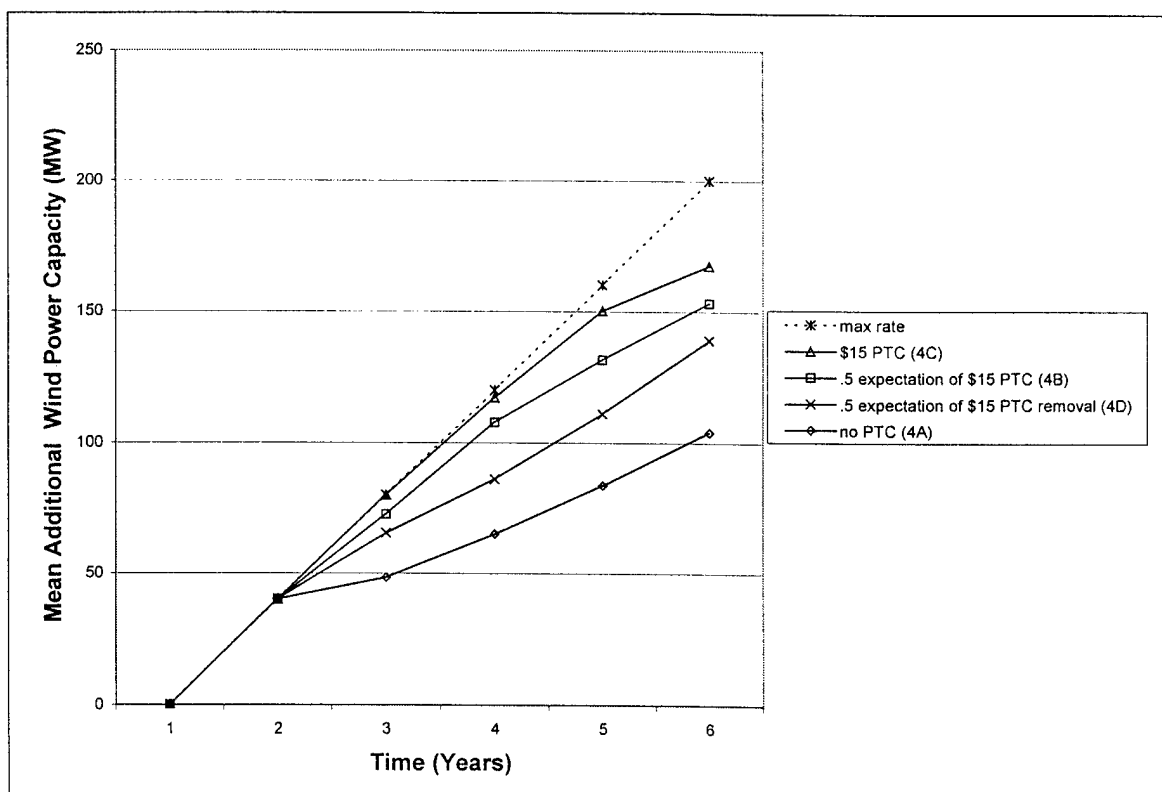


Figure 32. Expectation of PTC Removal or Addition (Scenario 4A – 4D)

Results from Scenarios 5 and 6 further elucidate these effects for the ITC and PTC as well as demonstrate the interaction between  $\lambda_0$  and  $\lambda_1$ . Results from Scenario 5 are shown in Tables 10 and 11 for states where the ITC is not in effect and is in effect respectively. Similarly, results from Scenario 6 are shown in Table 12 for all states where the PTC is not effect and in Table 13 for states where the PTC is in effect. In all cases cell values, which are rounded to the nearest integer, represent the mean investment level from the optimal RL policy across all states with an identical policy parameter. For instance, Table 10 cell values are averaged across all states in the state space where the ITC policy is in effect. The metric utilized in these tables has little meaning in absolute terms because it is highly dependent on the bounds of the state space. However, the metric is useful in relative terms because all state spaces are defined using identical upper bounds.

In Table 10, investment in wind power remains constant when  $\lambda_0$  is equal to zero and  $\lambda_1$  varies from zero to 0.5. This is expected because if the agent is in a state without the ITC and the probability of transitioning out of this state is zero, the probability of transitioning from an ITC state to a non-ITC state should not affect the firm's behavior. There is a sharp decrease in investment as  $\lambda_0$  increases above zero, due to the increased option value of postponing investment until the agent reaches an ITC state. There is also a slight interaction between  $\lambda_0$  and  $\lambda_1$ . For higher levels of  $\lambda_0$ , investment increases as  $\lambda_1$  increases because firms will postpone a larger share of their investments in wind power for a permanent change in policy than for a transient change. This reaction results from

the capacity restrictions that limit total wind power investment each period. If an ITC enactment is short-lived, a firm may not be able to invest as much as it wants while the policy is in place due to investment capacity restrictions. It is likely that this interaction would be less significant for higher upper bounds on per-period wind power investment.

Table 10. Mean Investment (MW) In Wind Power Across States without the ITC

$\lambda_0$	$\lambda_1$					
	0.0	0.1	0.2	0.3	0.4	0.5
0.0	20	20	20	20	20	20
0.1	16	17	18	18	18	18
0.2	14	15	15	16	16	17
0.3	13	14	14	15	15	16
0.4	13	13	14	14	15	15
0.5	12	13	13	13	14	14

In Table 11, there is no effect from increasing  $\lambda_0$  if  $\lambda_1$  is zero, because if there is no chance of leaving an ITC state, the probability of transitioning from a state without the ITC to a state with the ITC is irrelevant. As one looks across the table, the investment increasing effect of expecting transition to a state without the ITC can be seen. Firms increase their level of investment as the probability of ITC removal ( $\lambda_1$ ) increases. There is also an interaction between  $\lambda_1$  and  $\lambda_0$  at higher levels of  $\lambda_1$ . As  $\lambda_0$  increases, the investment-increasing effect of  $\lambda_1$  is mitigated. This occurs because if firms expect the removal of the ITC policy to be permanent, they are likely to invest at a higher level than if they believe that ITC removal will only be temporary.

Tables 10 and 11 illustrate that the investment increasing effect from the expectation of ITC removal is smaller than the investment inhibiting effect that results from the expectation of an ITC. In the case of an expected ITC addition (Table 10), the maximum investment decreasing effect is 40 percent (20 to 12) for  $\lambda_0$  equal to 0.5 and  $\lambda_1$  equal to 0.0. In contrast, in the case of the expected ITC removal, the maximum investment increasing effect is 13 percent (25 to 28) for  $\lambda_0$  equal to 0.0 and  $\lambda_1$  equal to 0.5. It is possible that the differences between the effect of an addition versus a removal would not be as great if the upper bound on wind power investment were relaxed. Given the current upper bound, it is possible that the investment increasing effect from a pending ITC removal is diminished by the upper bound on investment in wind power.

Table 11. Mean Investment (MW) in Wind Power across States with the ITC

$\lambda_0$	$\lambda_1$					
	0.0	0.1	0.2	0.3	0.4	0.5
0.0	25	27	28	28	28	28
0.1	25	26	27	27	28	28
0.2	25	26	26	27	27	27
0.3	25	26	26	26	27	27
0.4	25	26	26	26	26	27
0.5	25	26	26	26	26	27

Tables 12 and 13 show results for the PTC from those states without and with the PTC in place respectively. In Table 12, there are constant levels of investment when  $\lambda_0$  is zero due to the irrelevance of  $\lambda_1$  when there is no chance of transitioning from a state without the PTC to a state with the PTC. As  $\lambda_0$  increases, investment increases because of increased expectations that a PTC will be implemented in the future. This investment-

increasing effect is stronger at lower levels of  $\lambda_1$  because they increase the expected time until an enacted PTC is removed as well as increase the probability of the system being in a PTC state at all future points in time.

Table 12. Mean Wind Power Investment (MW) across States without the PTC

$\lambda_0$	$\lambda_1$					
	0.0	0.1	0.2	0.3	0.4	0.5
0.0	20	20	20	20	20	20
0.1	26	26	26	25	25	24
0.2	28	28	27	27	27	26
0.3	30	29	29	28	28	27
0.4	30	30	30	29	29	28
0.5	31	31	30	30	29	29

Table 13 shows mean wind power investment across states in which the PTC is in effect. As the probability of the PTC being removed ( $\lambda_1$ ) increases, investment is inhibited. Higher levels of  $\lambda_0$  mitigate this effect due to the influence of  $\lambda_0$  on the mean time until the agent transitions back to a state with the PTC and the probability of the agent being in a state with the PTC at all future time periods.

While the results from PTC are opposite in sign to those of the ITC, similar differences concerning the magnitudes of the effects of potential policy addition or removal can be seen. Table 12 shows the maximum amount of investment increase from expectation of the PTC of 51 percent (20 to 31) when  $\lambda_0$  is equal to 0.5 and  $\lambda_1$  is equal to 0.0. This contrasts with the maximum amount of investment decrease from an expected PTC removal of 21 percent (32 to 26) when  $\lambda_0$  is equal to 0.0 and  $\lambda_1$  is equal to 0.5.

Table 13. Mean Wind Power Investment (MW) across States with the PTC

$\lambda_0$	$\lambda_1$					
	0.0	0.1	0.2	0.3	0.4	0.5
0.0	32	31	30	28	27	26
0.1	32	31	31	29	28	27
0.2	32	31	31	30	29	28
0.3	32	32	31	30	29	29
0.4	32	32	31	30	29	29
0.5	32	32	31	30	30	29

#### 4.6 Conclusions and Policy Implications

This research demonstrates the strong relationship between policy uncertainty and investment behavior. We see that anticipation of a proposed policy change may produce near-term investment results that are opposite in direction to the intended result of the proposed change. Investment Tax Credits are one example of a policy that produces this reverse outcome because their benefits are only realized on investments made during periods in which the ITC is active. This effect is not observed with the PTC, as it is modeled in this essay, because the benefits from the PTC in a given year do not depend upon the year in which an initial investment is made. Rather, the benefits from the PTC are only determined by the policy in place in any given year.

Therefore, if legislation were introduced in Congress to provide a large ITC on wind power, investment may subside as firms wait for the credit to be enacted. These results show that even a very low likelihood of actual ITC enactment could motivate a large decrease in wind power investment. Similarly, uncertainty over whether a given

policy will be extended beyond its expiration date could speed up investment in wind power investment beyond desired levels. These effects contrast with the impact of expectations of a PTC, as it is modeled in this essay, which may produce an increase (decrease) in wind power investment prior to enactment (removal) of the PTC.

The results from this essay extend the work of Dixit and Pindyk (1994) and Hasset and Metcalf (1999) by looking at a case where ITCs are only applied to a subset of available technologies. Since, substitution opportunities exist between wind and classical technology investments, the investment postponing and enhancing effects of ITC expectation are stronger than those previously found.

Results from this essay also make clear that long-run policy stability is critical to effective management of wind power subsidy programs. However, since this is often impossible given the political nature of public policy in the United States, policies should be structured to provide benefits during the years in which the policy is in effect regardless of the year of investment.

The PTC that is currently in place in the United States does not operate in the same way as the PTC that is modeled in this essay. Rather, the PTC's benefits are only realized on investments made when the policy is in effect. Therefore, the expectation of this PTC's addition or removal would impact investment in a manner similar to that of an ITC. This effect may partially explain the large increase in wind power investment in 1998 and 1999 as investors increased their rate of investment to take advantage of the PTC before the extension/removal decision was made in 1999.

One policy recommendation stemming from this essay's results is that future PTCs should provide a multi-year guarantee of tax credits to all firms who invest during the period in which the legislation is in effect. However, these results suggest that a stipulation should be added allowing for firms to take advantage of the credit, while the policy is in effect, regardless of when their investment was made. A policy structured in this manner would be less prone to the strong investment decreasing effect of policy expectation than a policy that only rewards firms that invest while the policy is in effect, because firms that make their investment decision prior to policy enactment would still be able to realize some of the benefits of the policy. However, a policy structured in this manner would not prevent "over investment" upon the expectation of policy removal.

Results from this essay also suggest that policies that are not stable across time may bring about suboptimal increases or decreases in the investment level. This type of policy stability is not being fostered by current United States legislative actions. Rather than either ending the PTC or granting a 5-year extension of the PTC through June 30, 2004, a compromise two and a half year extension through December 31, 2001 was reached. This short extension may create another flurry of wind power investment activity in 2001 as investors rush to invest before the PTC's potential removal.

These results also provide some preliminary insights on the effect of pending RPS legislation on current investment in renewable technologies. Since the RPS requires that a certain percentage of a firm's generation come from renewable power or that firms buy renewable power credits, anticipation of this standard should encourage firms to invest in

renewable technologies prior to enactment of the legislation. This result should occur because the proposed RPS does not differentiate between investments made prior to and during the period in which the legislation is in place. Therefore, the anticipation of a RPS creates no incentive for firms to postpone renewable investments prior to enactment of the RPS. The lag between the investment decision and investment completion also deters firms from waiting to see if a RPS will be in effect prior to investing. Renewable investment levels prior to the RPS should still be lower than those levels while the RPS is in effect. In addition, investment in renewable technologies should be lower for firms facing an uncertain RPS enactment compared with firms that face certain RPS enactment.

Additionally, Several extensions to this research are suggested. First, the relationship between uncertain tradable pollution permits and wind power investment should be analyzed since this is an alternative means to encourage wind power investment (and renewables more generally). If firms are forced to either limit their fossil fuel emissions or purchase tradable permits, it should encourage investment in wind and other renewables by reducing the relative cost of renewable generation. It would also be useful to ascertain the effects of uncertain wind power ITCs, wind power PTCs, and tradable pollution permits on classical dispatch. This analysis would show how uncertainty over the aforementioned policies would affect actual pollution levels. It is likely that the degree of substitutability between wind power and the classical generation technologies would greatly impact the amount by which emissions could be reduced through wind power investment. Also, sensitivity analysis on the bounds of the

action space, the level of the ITC or PTC, and the wind power capacity factor would provide a better understanding of the relationship between these assumptions and the investment response to tax policy uncertainty.

Finally, the model presented in this essay could be extended to examine a RPS rather than a single technology subsidy. An RPS may be preferable to the single technology subsidy addressed in this paper because substitution among renewable technologies is permitted. Therefore, the market determines the mix of renewable technologies. This contrasts with a single technology subsidy, which creates a bias in favor of the subsidized technology. Also, policies such as a RPS give producers the greatest flexibility because they permit them to either purchase renewable credits or invest directly in renewable power depending on which alternative is most cost effective.

## Chapter 5

## CONCLUSIONS

This research developed a reinforcement learning (RL)-based modeling framework for analyzing long-run electricity generation investment and applied it to several relevant policy issues. The framework analyzed the effect of capacity subsidies and price caps on investment level and spot prices. Additionally, the framework was used to determine the effect of an anticipated investment tax credit (ITC) or production tax credit (PTC) enactment or removal on wind power investment.

The first essay demonstrated that reinforcement learning (RL) can be used to develop flexible models of generation investment behavior under uncertainty. The flexible nature of this technique results from the fact that RL does not require the explicit definition of transition probabilities and thereby circumvents the “curse of modeling.” Instead, an optimal policy is derived through a trial and error interaction between an agent and its environment. When RL is used to model generation investment, several general conclusions regarding electricity generation investment and uncertainty are demonstrated. First, the large up-front investment costs and per-period fixed costs associated with generation investment cause expected value models to significantly overestimate generation investment levels when demand is uncertain. This bias results from the failure of expected value models to account for the opportunity cost of investing when there is

the option to wait for more information. Similarly, the results showed that an overestimation of the level of demand uncertainty will lead to predicted investment outcomes that fall short of actual levels. These modeling biases are critical for policy makers to understand because many planning models that are currently used to forecast future investment do not account for uncertainty. If these models overestimate future levels of investment, policy makers may be surprised when actual investment levels fall short of these predictions. This direction of modeling bias is especially problematic because insufficient levels of investment could result in system reliability problems if no mechanisms are in effect to promote a demand-side response to price.

The second essay exploited the flexibility of RL to show how the design of a restructured electricity market can impact long-run investment behavior and spot market prices. The results showed that capacity subsidies act to increase overall investment while reducing spot market price volatility. These benefits come at the expense of increasing average total electricity prices, where the total prices include both energy prices and capacity charges.

The results suggest that capacity subsidies, or closely related reserve requirements, may not be the most efficient policy alternative for ensuring generation adequacy because they implicitly assume that all consumers have similar risk preferences. Additionally, this mechanism assumes that all consumers value reliability equally. Therefore, a forward market combined with a system in which consumers can

self-select the level of reliability they desire should attain the benefits of a capacity subsidy in a more efficient manner.

The results also showed that price caps will not reduce average prices in the social welfare maximization scenario and may actually raise average electricity prices. In addition, they may reduce overall investment levels and result in welfare losses if the ISO is forced to shed load. Therefore, since social welfare maximization will approximate a competitive outcome, the results imply that price caps should be avoided in competitive markets. In contrast, for the monopoly producer, price caps produce an indeterminate effect on overall investment, and unequivocally lower average prices (which are otherwise unbounded). Therefore, price caps are necessary to prevent unlimited price markups. However, the ideal level of a monopolist's price cap is difficult to determine as a result of the bimodal response of investment to the price cap level. If the policy goal is to maximize the investment level, then price caps ranging from \$200/MWh to \$300/MWh appear to be ideal. Lower caps may reduce investment levels because if prices are too low, the monopolist will not invest to meet peaking loads. Also, if the cap is too high, there will be a decrease in the investment level because the monopolist will reduce its level of output in all periods in order to increase the market spot price to the price cap level.

The third essay used RL to demonstrate how uncertainty over the enactment or repeal of a wind power subsidy may affect wind power investment. If the policy in question only rewards firms if they invest while the policy is in effect, as is the case with

an investment tax credit (ITC), firms will speed up their investment decisions if they anticipate a policy repeal. If firms anticipate a policy enactment, they will slow down their rate of investment due to the increased option value of waiting rather than investing. In contrast, if the policy in question is applied without regard to the year of investment, as is the case with the production tax credits (PTCs) modeled in this research, the direction of the effects from uncertainty will change. For this type of policy, firms will speed up their level of investment in anticipation of a policy enactment and slow down their rate of investment in anticipation of a policy repeal.

The third essay also demonstrated that the effects of policy uncertainty may be stronger when substitution opportunities exist between subsidized and unsubsidized technologies, because firms may make compensating increases or decreases in investment in the nonsubsidized technology. Because of the significance of the effects of pending policy enactment or removal, policy makers should strive to attain policy stability. If this is impossible, due to political or other factors, then policies should be designed so that they are applied without regard to the year of investment.

Each of the essays in this dissertation demonstrated the importance of considering dynamics and uncertainty when analyzing the magnitude and direction of policy effects on investment. A static analysis of the effects of a price cap will always lead to a lower average price if the cap is binding or an unchanged price if the cap is nonbinding. The second essay demonstrated that dynamic analysis may produce the opposite conclusion due to the effect of price caps on long-run investment. The criticality of considering

uncertainty and dynamics is also demonstrated in the third essay which evaluates the effects of policy uncertainty on investment. The result that a potential ITC enactment (removal) could lead to a decrease (increase) in the respective level of investment is counterintuitive and would be impossible to model statically. Even a two-period model could not replicate this analysis because it could not simultaneously consider a firm's expectations concerning the probability of policy removal ( $\lambda_1$ ) and the probability of policy enactment ( $\lambda_0$ ).

This research has shown that RL is a useful tool that can effectively model the effects of various policy issues on electricity generation investment. Future work could apply the RL framework to analyze investment in transmission in addition to generation. Future work should also focus on the development of multi-agent models that are capable of examining cases of imperfect competition. These multi-agent models could potentially capture the game theoretic aspects of oligopolistic markets.

## REFERENCES CITED

- Aizenman, Joshua and Nancy P. Marion. 1993. Policy Uncertainty, Persistence and Growth. *Review of International Economics* 1, no. 2: 145-163.
- American Wind Energy Association (AWEA). 1999. *Colorado and Wyoming Wind Energy Projects*. American Wind Energy Association. Washington D.C. Available from [www.awea.org/projects](http://www.awea.org/projects).
- \_\_\_\_\_. 2000a. *Renewable Portfolio Standard (RPS) Progresses in States and Congress as an Integral Part of a Restructured Electric Industry*. American Wind Energy Association. Washington D.C. Available from [www.awea.org/policy/rpsbrief](http://www.awea.org/policy/rpsbrief).
- \_\_\_\_\_. 2000b. *Wind Energy and Climate Change: A Proposal for a Strategic Initiative*. American Wind Energy Association. Washington D.C. Available from [www.awea.org/policy/ccwp.html](http://www.awea.org/policy/ccwp.html).
- \_\_\_\_\_. 2000c. *Wind Energy: U.S. Installed Capacity*. American Wind Energy Association. Washington D.C. Available from [www.awea.org/faq/instcap.html](http://www.awea.org/faq/instcap.html).
- Averch, H. and L. Johnson. 1962. Behavior of the Firm Under Regulatory Constraint. *American Economic Review* 52: 1053-1069.
- Awerbuch, Shimon. 2000. Getting it Right: The Real Cost Impacts of a Renewables Portfolio Standard. *Public Utilities Fortnightly* (Feb): 44-55.
- Barto, A.G., S.J. Bradtke, and S.P. Singh. 1991. Learning to Act Using Real-time Dynamic Programming. *Artificial Intelligence* 72: 81-138.
- Benner, J.H. 1992. *Impact of Wind Turbines on Birdlife: An Overview of Existing Data and Lacks in Knowledge in Order of the European Community*. Rotterdam, Netherlands: Consultants on Energy and the Environment.
- Bernow, Stephen, Bruce Biewald, and Donald Marron. 1991. Full-Cost Dispatch: Incorporating Environmental Externalities in Electric System Operation. *The Electricity Journal* 4, no. 2: 20-34.

- Bertsekas, D.P. and Tsitsiklis. 1996. *Neuro-Dynamic Programming*. Belmont, Mass.: Athena Scientific.
- Bhavaraju, Murty. November 1999. *Determination of Pool Forecast Requirement*. PJM ISO. Norristown, Penn. Available from [www.pjm.com](http://www.pjm.com).
- Birge, John R. and Francois Louveaux. 1997. *Introduction to Stochastic Programming*. New York: Springer.
- Bizer, D. and K. Judd. 1989. Taxation and Uncertainty. *American Economic Review* 79: 331-336.
- Booth, P.R. 1972. Optimal Generation Planning Considering Uncertainty. *IEEE Transactions on Power Apparatus and Systems* 91: 70-77.
- Borenstein, Severin. 1999a. Private Communication.
- \_\_\_\_\_. 1999b. *Understanding Competitive Pricing and Market Power in Wholesale Electricity Markets*. Berkeley: University of California Energy Institute.
- Borenstein, Severin and James Bushnell. 1998. *An Empirical Analysis of the Potential for Market Power in California's Electricity Industry*. Berkeley: University of California Energy Institute.
- Borenstein, Severin, James Bushnell, and Christopher Knittel. 1999. Market Power in Electricity Markets: Beyond Concentration Measures. *The Energy Journal* 20, no. 4: 65-88.
- Burtraw, Dallas, Karen Palmer, and Alan J. Krupnick. 1993. *Electric Utility Planning with Externalities*. Washington, D.C.: Resources for the Future.
- Byrne, Warren W. 1999. *Green Power in California*: Foresight Energy Company.
- Cavallo, A.J. 1995. High-Capacity Factor Wind Energy Systems. *Journal of Solar Energy* 117: 137-143.
- Conkling, Roger L. 1998. A California Generation Capacity Market. *The Electricity Journal*, (October): 27-35.
- Connel, J. and Mahadevan. 1993. *Robot Learning*. Boston: Kluwer Academic.
- Courville. 1974. Regulation and Efficiency in the Electric Utility Industry. *Bell Journal of Economics and Management Science* 5: 54-74.

- Cox, Alan J., Carl J. Blumstein, and Richard J. Gilbert. 1991. Wind Power in California: A Case Study of Targeted Subsidies. In *Regulatory Choices: A Perspective on Developments in Energy Policy*, ed. Richard J. Gilbert. Berkeley: University of California Press.
- Crandall, Robert and Jerry Ellig. 1997. *Economic Deregulation and Customer Choice: Lessons for the Electric Industry*. Fairfax, Va: Mercatus Center, George Mason University.
- Crites, Robert and Andrew Barto. 1996. *Advances in Neural Information Processing*. Cambridge: MIT Press.
- Desvousges, William H., F. Reed Johnson, and H. Spencer Banzhaf. 1994. *Externality Costs for Electricity Generation*. Durham, N.C.: Triangle Economic Research.
- Dixit, Avinash K. and Robert S. Pindyk. 1994. *Investment Under Uncertainty*. Princeton: Princeton University Press.
- Department of Energy (DOE). 1997. *Scenarios of U.S. Carbon Reductions*. Washington, D.C.: United States Department of Energy.
- Department of Energy (DOE). 1998. *Maintaining Reliability in a Competitive U.S. Electricity Industry*. Washington, D.C.: U.S. Department of Energy.
- Erev, Ido and Alvin E. Roth. 1998. Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *The American Economic Review* 88, no. 4: 848-881.
- Fehr, Nils-Henrick M von der and David Cameron Harbord. 1997. *Capacity Investment and Competition in Decentralised Electricity Markets*. Oslo, Norway: University of Oslo.
- Fox-Penner, Peter. 1997. *Electric Utility Restructuring: A Guide to the Competitive Era*. Vienna, Va.: Public Utilities Reports, Inc.
- Freeman, Myrick and Robert Rowe. 1995. Ranking Electric Generating Technologies with External Costs. *The Electricity Journal*, (December): 51-57.
- Gal-Or, E. and M. Spiro. 1992. Regulatory Regimes in the Electric Power Industry: Implications for Capacity. *Journal of Regulatory Economics* 4: 263-278.
- Gipe, Paul. 1995. *Wind Energy Comes of Age*. New York: John Wiley & Sons.

- Graves, F.C., E.G. Read, P.Q. Hanser, and R.L. Earle. 1998. One-Part Markets for Electric Power: Ensuring the Benefits of Competition. Cambridge: The Brattle Group.
- Green, Richard and David M Newberry. 1992. Competition in the British Electricity Spot Market. *Journal of Political Economy* 100: 929-953.
- Gas Research Institute (GRI). 2000. *Combustion Turbine Design, Operation, and Control*. Gas Research Institute. Chicago. Available from [www.gri.org](http://www.gri.org).
- Hassett, Kevin A. and Gilbert E. Metcalf. 1999. Investment With Uncertain Tax Policy: Does Random Tax Policy Discourage Investment? *The Economic Journal* 109: 372-393.
- Henney, A. 1998. Contrasts in Restructuring Wholesale Electric Markets: England/Wales, California, and the PJM. *The Electricity Journal* 11, no. 7: 24-42.
- Herbelot, Olivier. 1992. Option Valuation of Flexible Investments: The Case of Environmental Investments in the Electric Power Industry. Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge.
- Hirst, Eric, Brendan Kirby, and Stan Hadley. 1999. *Generation and Transmission Adequacy In A Restructuring U.S. Electricity Industry*: Edison Electric Institute.
- Hogan, William. 1998. *Competitive Electricity Market Design: A Wholesale Primer*. Cambridge: JFK School of Government, Harvard University.
- Jaakula, S. Singh, and M. Jordan. 1994. On the Convergence of Stochastic Iterative Dynamic Programming Algorithms. *Neural Computation* 6, no. 6: 1185-1201.
- Jaffe, Adam and Frank Felder. 1996. Should Electricity Markets Have a Capacity Requirement? If So, How Should it Be Priced. *The Electricity Journal* (Dec): 52-60.
- Kittel, Charles. 1996. *Introduction to Solid State Physics*. New York: John Wiley and Sons.
- Krugman, Paul R. and Maurice Obstfel. 1992. *International Economics Theory and Policy*. Menlo Park, Calif.: Addison-Wesley.
- Leston, David. 1992. Investment Decisions and Transferable Discharge Permits: An Empirical Study of Water Quality Management under Policy Uncertainty. *Environmental and Resource Economics* 2: 441-458.

- Levin, N., A. Tishler, and J. Zahavi. 1983. Time Step vs Dynamic Optimization on Generation Capacity Expansion Programs of Power Systems. *Operations Research* 31, no. 5: 891-914.
- McDonald, Robert and Daniel Siegel. 1986. The Value of Waiting to Invest. *The Quarterly Journal of Economics* 101, no. 4: 331-349.
- Michaels, Robert J. 1997. MW Gamble: The Missing Market for Capacity. *The Electricity Journal* (December): 56-64.
- Mintzer, Irving M., Alan S. Miller, and Adam Serchuk. 1996. *The Environmental Imperative: A Driving Force in the Development and Deployment of Renewable Energy Technologies*. Washington D.C.: American Wind Energy Association.
- Moody, J. 1996. Optimization of Trading Systems and Portfolios. In *Proceedings of the Fourth International Conference on Neural Networks in Capital Markets*. Pasadena Calif.
- Nijkamp, Peter. 1977. *Theory and Application of Environmental Economics*. Edited by Ake Anderson and Walter Isard. Studies in Regional Science and Urban Economics. Amsterdam: Free University.
- Orloff, S. and A. Flannery. 1992. *Wind Turbine Effects on Avian Activity, Habitat Use, and Mortality in Altamont Pass and Solano County*. Sacramento: Biosystems Analysis Inc.
- Orr, Robert W. 1988. Irreversible Capacity Investment Under Uncertainty: An Application to Alberta's Electrical Generation Industry. M.A. Thesis, University of Calgary, Calgary, Canada.
- Palmer, Karen and Hadi Dowlatabadi. 1993. Implementing Environmental Costing in the Electric Utility Industry. *Energy and the Environment* 4, no. 3: 197-220.
- Pindyck, Robert S. 1991. Irreversibility, Uncertainty, and Investment. *Journal of Economic Literature* 29: 1110-1148.
- Pennsylvania-New Jersey-Maryland Interconnection (PJM). 1998a. *Reliability Assurance Agreement Among Load Serving Entities in the PJM Control Area*. Norristown, Penn.: PJM Interconnection.
- \_\_\_\_\_. 1998b. *Schedule 11: PJM Capacity Credit Market*. Norristown, Penn.: PJM Interconnection.

- \_\_\_\_\_. 2000c. *PJM Home Page*. Pennsylvania-New Jersey-Maryland Interconnection. Norristown, Penn. Available from [www.pjm.com](http://www.pjm.com).
- \_\_\_\_\_. 2000d. *Capacity Credit Market Operations*. Pennsylvania-New Jersey-Maryland Interconnection. Norristown, Penn. Available from [www.pjm.com](http://www.pjm.com).
- Poulsen, Johannes. 1999. Vestas and the US Market in the New Millennium. In *World Energy Congress*. Houston, Texas.
- Putnam, Robert. 1996. How Difficult is it to Integrate Wind Turbines With Utilities. *Wind Energy Weekly* (January): 680-686.
- Quick, David M. 2000. The Effect Transmission on a Dominant Firm's Ability to Exert Market Power in the Electricity Industry. Ph.D. Dissertation, Colorado School of Mines, Golden, Colo.
- Righter, Robert W. 1996. Wind Power. In *Wind Energy In America*:361. Norman, Okla: University of Oklahoma Press.
- Ross, Sheldon. 1982. *Introduction to Stochastic Dynamic Programming*. Orlando: Academic Press.
- Roth, Alvin E. and Ido Erev. 1995. Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior* 8, no. 1: 164-212.
- Roth, Alvin E., Vesna Prasnikar, Masashiro Okuno-Fujiwara, and Shmuel Zamir. 1991. Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study. *American Economic Review* 81, no. 5: 1068-1095.
- Rowe, Robert D. , Steve Bernow, and David White. 1995. *New York State Environmental Externalities Cost Study*. Boulder, Colo.: RCG/Haglar Bailly Inc.
- Ruff, Larry E. 1999. *Competitive Electricity Markets: Why They are Working And How to Improve Them*: National Economic Research Associates, Washington D.C.
- Rust, John, ed. 1996a. *Numerical Dynamic Programming in Economics*. Edited by H.A. Amman, D.A. Kendrick, and J. Rust. Handbook of Computational Economics. New York: Elsevier.
- \_\_\_\_\_. 1996b. Using Randomization to Break the Curse of Dimensionality. *Econometrica* 65, no. 3: 487-516.

- Sherali, A.L., F.H. Soyster, and Murph F.H. 1984. Intertemporal Allocation of Capital Costs in Electric Utility Capacity Expansion Planning Under Uncertainty. *Management Science* 30, no. 1: 1-19.
- Singh, Harry and Jonathan Jacobs. 2000. Capacity Products and ISO Markets. In *Electricity Industry Restructuring: Fifth Annual Research Conference*. Berkeley: University of California.
- Singh, S. and D. Bertsekas. 1997. Reinforcement Learning for Dynamic Channel Allocation in Cellular Telephone Systems. In *Advances in Neural Information Processing Systems*: 974-980. Cambridge: MIT Press.
- Smith, James E. and Kevin F. McCardle. 1999. Options in the Real World: Lessons Learned in Evaluating Oil and Gas Investments. *Operations Research* 47, no. 1: 1-15.
- Stermole, Franklin J. and John M. Stermole. 1996. *Economic Evaluation and Investment Decision Methods*. Golden, Colo.: Investment Evaluation Corporation.
- Steve, Jaime. 1999. *The Production Tax Credit*. Washington D.C.: House of Representatives Committee on Ways and Means.
- Stoll, Harry H. 1989. *Least-Cost Utility Planning*. New York: John Wiley & Sons.
- Stone and Webster Management Consultants. 1998. *Energy and Economic Modeling Issues Related to an Evaluation of the Regulatory Structure of the Retail Electric Industry in the State of Colorado*. Englewood, Colo.: Colorado Electric Advisory Panel.
- Stuebi, Richard. 1999. *The Rise, Fall and Rebirth of Wind Energy in California: Lessons for the Renewable Energy Industry*. Denver: NextWaveEnergy.
- Sulkko, Andy. 1999. Private Communication
- Sutton, Richard and Andrew Barto. 1998. *Reinforcement Learning*. Cambridge: MIT Press.
- Sweetser, Al. 1998. An Empirical Analysis of a Dominant Firm's Market Power in a Restructured Electricity Market, A Case Study of Colorado. Ph.D. Dissertation, Colorado School of Mines, Golden, CO.
- Teisberg-Olmsted, Elizabeth. 1993. Capital Investment Strategies Under Uncertain Regulation. *RAND Journal of Economics* 24, no. 4: 591-604.

- Tesauro, Gerald. 1995. TD-Gammon, A Self Teaching Backgammon Program, Achieves Master Level Play. *Neural Computation* 9, no. 2: 215-219.
- Van Roy, B. 1998. Learning and Value Function Approximation in Complex Decision Processes. Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge.
- Wang, Earl-Juei, Majid Jaraiedi, and Thomas F. Torries. 1996. Modeling Long-Run Cost Minimization and Environmental Provisions for Utility Expansion. *Energy Economics* 18: 49-68.
- Watkins, C.J. 1989. Learning From Delayed Rewards. Ph.D. Dissertation, Cambridge University, Cambridge, England.
- Weitzman, Martin. 1974. Prices vs. Quantities. *Review of Economic Studies* 41: 477-491.
- Winston, Wayne L. 1994. *Operations Research: Applications and Algorithms*. Belmont, Calif.: Duxbury Press.
- Wiser, Ryan, Kevin Porter, and Steve Clemmer. 1999. Emerging Markets for Wind Power: The Role of State Policies Under Restructuring. In *WINDPOWER 1999*. Burlington, Vt.: AWEA.
- Wolak, Frank and Robert H. Patrick. 1996. *The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market*. Palo Alto, Calif.: Stanford University.
- Wolak, Frank A., Roger E. Bohn, James B. Bushnell, and Alvin K. Klevorick. 1999. *Memorandum On Maximum Purchase Prices in Energy and Ancillary Services Markets in California*. Folsom, CA: California Independent System Operator.
- Zajac, E. 1970. A Geometric Treatment of Averch-Johnson's Behavior of the Firm Model. *American Economic Review* 79: 331-336.

## APPENDIX A

## C++ CODE FOR GENERAL RL FRAMEWORK

```

#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <iostream.h>
//technology 1 is combined cycle
//technology 2 is combustion turbine

//*****CONSTANTS*****
#define maxk1 19 //
#define maxk2 19//
#define maxd 19//
#define numstates 8000// $(\text{maxk1}+1)*(\text{maxk2}+1)*(\text{maxd}+1)$ 
#define maxact 6 //total number of actions available at any time step
#define epsilon .75
#define theta 1
#define alpha 40 //slope of linear demand curve
#define gamma .9 //1/(1+discount rate)
#define v1 17
#define v2 26
#define f1 11110
#define f2 150
#define i1 573000//combined cycle
#define i2 384000//combustion turbine
#define maxloads 8
#define simtime 60//years in simulation
#define simnum 100//replications of simulation
#define block 150 //block size
#define dblock 1//additional demand scaling for block
#define instep 1//how many blocks you move each investment
#define startk1 0
#define startk2 0
#define startdemand 2
#define avail .9//plant availability

```

```

//*****GLOBAL VARIABLES
int k1,kk1,k2,kk2,d,dd; //capacity for each technology and demand shift
int s,ss;//current state
int a;//action chosen
double Q[numstates][maxact];
double soft[numstates][maxact];
double rewg[maxk1+1][maxk2+1][maxd+1][maxact];
int f[numstates];
int count;
int aa; //successor action
double loadsize[maxloads]; //load duration curve sizes
double load[maxloads]; //load duration curve loads
double sigma; //standard deviation of demand shift paramet transition
int perspective; //1=SW max; 0=profit max
int perspective2; //0=nothing 1=price adder 2=capacity payment
double pricecap;
double elas;
double sumdelta;
double sumdeltaold; //old sumdelta
double Qold;
double Qnew;
double lr; //learning rate
double first; //first derivative;
int showprice;
double meanprice; //mean price to be used to simulation;
double nondispatch; //amount of undispached capacity at peak
double res; //undispached capacity as a percentage
double peakprice;
double p0price;
double p1price;
double p2price;
double p3price;
double p4price;
double p5price;
double p6price;
double p7price;
double csg;
double prof;
double temperature0;
double temperature;

```



```

    }
    else if (sigma==1)
    {
        r=(rand()%100);
        z=looktable1[r];
    }
    else if (sigma==2)
    {
        r=(rand()%100);
        z=looktable2[r];
    }
    return(z);
}

//*****GETZ*****
*****
//*****SHOWMAX*****
*****
void showmax ()
{
    int i,j;//counters
    double temp;
    int maxstate;

    maxstate=0;
    temp=0;
    for (i=0;i<numstates;i++)
    {
        for (j=0;j<maxact;j++)
        {
            if (soft[i][j]>temp)
            {
                temp=soft[i][j];
                maxstate=i;
            }
        }
    }

    printf(" %i temp = %f probabilities = %f %f %f %f\n",maxstate,temperature,soft[maxstate][0],soft[maxstate][1],soft[maxstate][2],soft[maxstate][3]);

```

```

}
//*****SHOWMAX*****
*****
//*****INITIALIZE*****
*****
void initialize ()
{
    int i,j; //counters

    for (i=0;i<numstates;i++)
    {
        f[i]=0;
        for (j=0;j<maxact;j++)
        {
            Q[i][j]=1;
        }
    }

}
//*****INITIALIZE*****
*****
//*****INITDSHIFT*****
*****
void initdshift()
{
    int i,t; //counters
    double d;

    for (i=0;i<simnum;i++)
    {
        d=startdemand;
        for (t=0;t<simtime;t++)
        {
            dshiftvec[i][t]=d;
            d=d+theta+getz(sigma);
            if (d<0){d=0;}
            if (d>maxd){d=maxd;}
        }
    }
}

```

```

}
//*****INITDSHIFT*****
*****
//*****SOFTMAX*****
*****
void softmax ()
{
    int i,j;
    double total;
    double temp[maxact];
    double max;
    int t;

    total=0;

    for (i=0;i<numstates;i++)
    {
        max=-999999999;
        for (j=0;j<maxact;j++)
        {
            if (Q[i][j]>max)
            {
                max=Q[i][j];
                t=j;
            }
        }

        for (j=0;j<maxact;j++)
            {temp[j]=Q[i][j]/max;}

        total=0;
        for (j=0;j<maxact;j++)
        {
            total=total+pow(2.78,temp[j]/temperature);
        }

        for (j=0;j<maxact;j++)
        {
            soft[i][j]=(pow(2.78,temp[j]/temperature)/total);

```

```

    }

}

}
//*****SOFTMAX*****
*****
//*****SOFTACTION*****
*****
void softaction()
{
    double temp;
    int iii;
    double r;
    double toggle;

    r=((double)rand()/((double)RAND_MAX));//generates random number between 0
and 1

    toggle=0;
    temp=0;

    for (iii=0;((iii<maxact)&&(toggle==0));iii++)
    {
        temp=temp+soft[s][iii];
        if ((temp>r)&&(toggle==0))
        {
            a=iii;
            toggle=1;
        }
    }

}
//*****SOFTMAX*****
*****
//*****FINDSS*****
*****
void findss ()
{
    kk1=k1+instep*tech1[a];
    kk2=k2+instep*tech2[a];

```

```

dd=d+theta+getz(sigma);

if (kk1>maxk1) {kk1=maxk1;}
if (kk2>maxk2) {kk2=maxk2;}
if (kk1<0) {kk1=0;}
if (kk2<0) {kk2=0;}
if (dd>maxd) {dd=maxd;}
if (dd<0) {dd=0;}
}
//*****FINDSS*****
*****
//*****GETREWARD*****
*****
double getreward (int k1,int k2,int d,int a)
{

double qu1,qu2,q;//unconstrained dispatch of each technology
double q1,q2;//actual dispatch of each technology
double reward,totalreward;
double cap1,cap2,acap1,acap2;
double dtotal;//total demand parameter
int ii;//counter
double k10,k20;//initial quantities of capacity
double price,chokeprice,chokequantity;//price in this period
double cs;//consumer surplus
double temp2;//used in anchor point calculation

loadsize[0]=.00014;
loadsize[1]=.003871;
loadsize[2]=.202869;
loadsize[3]=.244422;
loadsize[4]=.318761;
loadsize[5]=.174863;
loadsize[6]=.050091;
loadsize[7]=.005009;

temp2=pow(30,elas);

load[0]=4000/temp2;
load[1]=5000/temp2;

```

```

load[2]=6000/temp2;
load[3]=7000/temp2;
load[4]=8000/temp2;
load[5]=9000/temp2;
load[6]=10000/temp2;
load[7]=11000/temp2;

chokeprice=1000;

meanprice=0;
quantity=0;

totalreward=0;

if (showprice==1)
    {csg=0;
    prof=0;}

k10=10000;
k20=0;

cap1=k1*block+k10;
cap2=k2*block+k20;

acap1=cap1*avail;
acap2=cap2*avail;

for (ii=0;ii<maxloads;ii++)
{

    dtotal=load[ii]+d*block*dblock;

    chokequantity=dtotal*pow(chokeprice,elas);

    if (perspective==1)
    {
        qu1=dtotal*pow(v1,elas);
        qu2=dtotal*pow(v2,elas);
    }
    else if (perspective==2)

```

```

    {
        qu1=dtotal*pow(pricecap,elas);
        qu2=dtotal*pow(pricecap,elas);
    }

    if (perspective==1)//SW MAX
    {
        if (acap1>qu1) {q=qu1;q1=qu1;q2=0;}
        else if ((acap1<qu2) && ((acap1+cap2)>qu2))
        {q=qu2;q1=acap1;q2=(qu2-acap1);}
        else if ((acap1>qu2) && (acap1<qu1))
        {q=acap1;q1=acap1;q2=0;}
        else {q=acap1+acap2;q1=acap1;q2=acap2;}
    } else if (perspective==2)//monopoly profit max
    {
        if (acap1>qu1) {q=qu1;q1=qu1;q2=0;}
        else if ((acap1<qu2) && ((acap1+acap2)>qu2))
        {q=qu2;q1=acap1;q2=(qu2-acap1);}
        else if ((acap1>qu2) && (acap1<qu1))
        {q=acap1;q1=acap1;q2=0;}
        else {q=acap1+acap2;q1=acap1;q2=acap2;}
    }

    if (q<chokequantity) {price=1000;}
    else {price=pow((q/dtotal),(1/elas));} //pow is still here

    if (price>pricecap) {price=pricecap;}

    if (showprice==1)
    {
        meanprice=meanprice+(price*loadsize[ii]);
        quantity=quantity+(q*loadsize[ii]);
    }

    cs=dtotal*pow(chokeprice,(elas+1))/(elas+1)-
        dtotal*pow(price,(elas+1))/(elas+1); //pow is still
here

    if (perspective==1)
    {
        reward=loadsize[ii]*365*24*

```

```

                                (q*price - v1*q1 - v2*q2 + cs);
    } else if (perspective==2)
    {
                                reward=loadsize[ii]*365*24*
                                (q*price - v1*q1 - v2*q2);
    }

    totalreward=totalreward+reward;

    if (showprice==1)
        {csg=csg+cs*loadsize[ii]*365*24;
        prof=prof+loadsize[ii]*365*24*(q*price-v1*q1-v2*q2);}

    if ( (showprice==1) && (ii==(maxloads-1)) )
{nondispatch=cap1+cap2-q;

    res=(cap1+cap2-q)/q;

    peakprice=price;}

    if ( (showprice==1) && (ii==0) ) {p0price=price;}
    if ( (showprice==1) && (ii==1) ) {p1price=price;}
    if ( (showprice==1) && (ii==2) ) {p2price=price;}
    if ( (showprice==1) && (ii==3) ) {p3price=price;}
    if ( (showprice==1) && (ii==4) ) {p4price=price;}
    if ( (showprice==1) && (ii==5) ) {p5price=price;}
    if ( (showprice==1) && (ii==6) ) {p6price=price;}
    if ( (showprice==1) && (ii==7) ) {p7price=price;}

} //for loop

totalreward=totalreward-f1*cap1 - f2*cap2-
                                instep*block*tech1[a]*i1-instep*block*tech2[a]*i2;

```

```

        if (showprice==1) {prof=prof-f1*cap1-f2*cap2-
                           instep*block*tech1[a]*i1-instep*block*tech2[a]*i2;}
        return(totalreward);
    }
    /*******GETREWARD*****
    *****/

    /*******UPDATEP*****
    *****/
    void updatep ()
    {
        int i;
        double temp;
        double temp2;

        temp2=f[s];
        temp=-999999;
        for (i=0;i<maxact;i++)
        {
            if (Q[s][i]>temp)
            {
                temp=Q[s][i];
                f[s]=i;
            }
        }
        if (temp2!=f[s]) {numpchanges++;}
    }
    /*******UPDATEP*****
    *****/

    /*******GETSTATE*****
    *****/
    int getstate(int k1,int k2,int d)
    {
        int temp;
        temp=d*((maxk1+1)*(maxk2+1))+k2*(maxk1+1)+k1;
        return(temp);
    }
    /*******GETSTATE*****
    *****/

```

```

//*****ABSOLUTE*****
*****
double absolute(double t1,double t2)
{
    double temp;
    temp=t1-t2;
    if (temp<0) {temp=temp*-1;}
    return(temp);
}
//*****ABSOLUTE*****
*****

//*****SHOWRESULTS*****
*****
void showresults ()
{
    int i,j,k;//counters
    int temp;
    FILE *SP;
    SP= fopen("show.dat","a");

    for (i=0;i<=maxk1;i++)
    {
        printf(" k1 = %i \n",i);
        fprintf(SP," k1 = %i \n",i);
        for (j=0;j<maxk2;j++)
        {
            for (k=0;k<maxd;k++)
            {
                temp=getstate(i,j,k);
                printf(" %i ",f[temp]);
                fprintf(SP," %i ",f[temp]);
            }
            printf(" \n");
            fprintf(SP," \n");
        }
    }
    fclose(SP);
}

```

```

//*****SHOWRESULTS*****
*****
//*****PRINTOUT*****
*****
void printout (double invar[simnum][simtime],int all)
{
    FILE *FP;
    FP= fopen("diss.dat","a");
    int i,t;//counters
    double mean[simtime];
    double stddev[simtime];
    double upper[simtime];
    double lower[simtime];
    //calcalating means
    for (t=0;t<simtime;t++)
    {
        mean[t]=0;
        for (i=0;i<simnum;i++)
        {
            mean[t]=mean[t]+(double)invar[i][t];
        }
        mean[t]=(mean[t]/simnum);
    }

    //calculating stddev, upper, and lower 95% confidence bounds
    for (t=0;t<simtime;t++)
    {
        stddev[t]=0;
        for (i=0;i<simnum;i++)
        {
            stddev[t]=stddev[t]+(mean[t]-invar[i][t])*(mean[t]-
invar[i][t]);
        }
        stddev[t]=(stddev[t]/(simnum-1));
        stddev[t]=pow(stddev[t],.5);
        lower[t]=mean[t]-1.96*stddev[t];
        upper[t]=mean[t]+1.96*stddev[t];
    }

    for (t=0;t<simtime;t++)

```

```

    {
        printf(" %f ",mean[t]);
        fprintf(FP," %f ",mean[t]);
    }
    printf("\n");
    fprintf(FP,"\n");

    if (all==1)
    {
        for (t=0;t<simtime;t++)
        {
            printf(" %f ",lower[t]);
            fprintf(FP," %f ",lower[t]);
        }
        printf("\n");
        fprintf(FP,"\n");

        for (t=0;t<simtime;t++)
        {
            printf(" %f ",upper[t]);
            fprintf(FP," %f ",upper[t]);
        }
        printf("\n");
        fprintf(FP,"\n");
    }//all
    fclose(FP);
}
//*****PRINTOUT*****
*****

//*****SIMULATE*****
*****

void simulate ()
{
    int i,t; //counters
    int ktech1,ktech2,demand;
    int kktech1,kktech2,ddemand;
    int simstate;//simulation state
    int sucstate;//successor state
    double reward;

```

```

double capacity[simnum][simtime]; //capacity by time and simulation run
double capacity1[simnum][simtime]; //capacity of technology 1
double capacity2[simnum][simtime]; //capacity of technology 2
double pricevec[simnum][simtime]; //price
double quantityvec[simnum][simtime]; //mean quantity dispatched
double nondispatchvec[simnum][simtime]; //amount of excess capacity at peak
double resvec[simnum][simtime]; //reserve margin
double peakpricevec[simnum][simtime];
double p0vec[simnum][simtime];
double p1vec[simnum][simtime];
double p2vec[simnum][simtime];
double p3vec[simnum][simtime];
double p4vec[simnum][simtime];
double p5vec[simnum][simtime];
double p6vec[simnum][simtime];
double p7vec[simnum][simtime];
double csvec[simnum][simtime];
double profvec[simnum][simtime];
double dshiftvec2[simnum][simtime];
double lolpvec[simnum][simtime];

```

```

for (i=0;i<simnum;i++)
{
    ktech1=startk1; //initial starting state parameters
    ktech2=startk2; //initial starting state parameters
    demand=startdemand; //initial starting state parameters
    for (t=0;t<simtime;t++)
    {
        capacity[i][t]=(ktech1+ktech2)*block;
        capacity1[i][t]=ktech1*block;
        capacity2[i][t]=ktech2*block;
        simstate=getstate(ktech1,ktech2,demand);
        kktech1=ktech1+instep*tech1[(f[simstate])];
        kktech2=ktech2+instep*tech2[(f[simstate])];
        ddemand=(int)dshiftvec[i][t+1];

        if (kktech1>maxk1) {kktech1=maxk1;}
        if (kktech2>maxk2) {kktech2=maxk2;}
        if (kktech1<0) {kktech1=0;}
    }
}

```

```

        if (kktech2<0) {kktech2=0;}

        showprice=1;
        reward=getreward(ktech1,ktech2,demand,f[simstate]);

        pricevec[i][t]=meanprice;//note: meanprice was calculated in
getreward

        quantityvec[i][t]=quantity;
        nondispatchvec[i][t]=nondispatch;//nondispatch was calculated in
get reward

        resvec[i][t]=res;//res was calculated in get reward
        peakpricevec[i][t]=peakprice;
        lolpvec[i][t]=lolp;
        p0vec[i][t]=p0price;
        p1vec[i][t]=p1price;
        p2vec[i][t]=p2price;
        p3vec[i][t]=p3price;
        p4vec[i][t]=p4price;
        p5vec[i][t]=p5price;
        p6vec[i][t]=p6price;
        p7vec[i][t]=p7price;

        csvec[i][t]=csg;

        profvec[i][t]=prof;

        dshiftvec2[i][t]=demand;

        showprice=0;

        sucstate=getstate(kktech1,kktech2,ddemand);
        simstate=sucstate;
        ktech1=kktech1;
        ktech2=kktech2;
        demand=ddemand;

        }//time loop
    }//i loop

    printf("total capacity ");

```

```
printout(capacity,1);
printf("technology 1 ");
printout(capacity1,1);
printf("technology 2 ");
printout(capacity2,1);
printf("price ");
printout(pricevec,0);
printf("quantity ");
printout(quantityvec,0);
printf("nondispatched capacity at peak ");
printout(nondispatchvec,0);
printf("reserve margin ");
printout(resvec,0);
printf("peak price ");
printout(peakpricevec,0);
printf("price 0 ");
printout(p0vec,0);
printf("price 1 ");
printout(p1vec,0);
printf("price 2 ");
printout(p2vec,0);
printf("price 3 ");
printout(p3vec,0);
printf("price 4 ");
printout(p4vec,0);
printf("price 5 ");
printout(p5vec,0);
printf("price 6 ");
printout(p6vec,0);
printf("price 7 ");

printf("cs ");
printout(csvec,0);

printf("prof ");
printout(profvec,0);

printf("demand ");
printout(dshiftvec2,1);
```

```

printf("lolp ");
printout(lolpvec,0);

} //simulate
//*****SIMULATE*****
*****

//*****INITREWG*****
*****
void initrewg()
{
    int i,j,k,m;

    for (i=0;i<=maxk1;i++)
    {
        for (j=0;j<=maxk2;j++)
        {
            for (k=0;k<=maxd;k++)
            {
                for (m=0;m<maxact;m++)
                {
                    rewg[i][j][k][m]=getreward(i,j,k,m);
                }
            }
        }
    }
}
//*****INITREWG*****
*****

//*****MYMAIN*****
*****
void mymain ()
{ //mymain
    FILE *PP;

    PP= fopen("detail.dat","a");

    srand(1);
    // lr=.4;

```

```

lr=.5;
harmonic=2;//3000
showprice=0;
numpchanges=0;
nopchanges=0;
temperature=1;
con=-1*log(.003)/maxcounta;//for temperature decay
con2=-1*log(.0001)/maxcounta;
initialize();//initialize all variables
initrewg();
softmax();
showmax();
for (count=0;nopchanges<T;count++)
{
    if (count==100000) {sumdeltaold=sumdelta;}
    if (((count%100000)==0)&&(count>100000))//1000 is baseline lr
decay=.999 is also baseline
    {
        first=sumdelta-sumdeltaold;
        if (first>0) {lr=lr*.1;}
        sumdeltaold=sumdelta;
        if ((count%100000)==0) {printf("it = %i numpchanges = %f
nopchanges = %f lr = %g sumdelta = %g
\n",count,numpchanges,nopchanges,lr,sumdelta);}
        sumdelta=0;
        if (numpchanges==0) {nopchanges++;}
        else {nopchanges=0;}
        numpchanges=0;
    }

    if (((count%100000)==0)&&(count>0))
    {
        temperature=(pow(2.78,-1*con*count));
        softmax();
        showmax();
        showresults();
    }

    for (k1=0;k1<=maxk1;k1++)
    {

```

```

        for (k2=0;k2<=maxk2;k2++)
        {
            for (d=0;d<=maxd;d++)
            {
                s=getstate(k1,k2,d);
                softaction();
                findss();
                ss=getstate(kk1,kk2,dd);
                aa=f[ss];
                Qold=Q[s][a];
                Q[s][a]=Q[s][a]
+lr*(rewg[k1][k2][d][a]+gamma*Q[ss][aa]-Q[s][a]);
                Qnew=Q[s][a];
                sumdelta=sumdelta+absolute(Qnew,Qold);
                updatep();
            } //d loops
        } //k2 loop
    } //k1 loop
} //count loop
printf("\n");
showprice=1;
simulate();
showresults();
FILE *FP;
FP= fopen("diss.dat","a");
fprintf(FP,"total number of iterations = %i sigma = %f\n",count,sigma);
fclose(FP);
fclose(PP);

} //mymain
//*****MYMAIN*****
*****
//*****MAIN*****
*****
void main ()
{ //main

    FILE *FP;
    FP= fopen("diss.dat","w");
    fclose(FP);

```

```

FILE *SP;
SP= fopen("show.dat","w");
fclose(SP);

FILE *PP;
PP= fopen("detail.dat","w");
fclose(PP);

sigma=1;
initdshift();

sigma=1;
maxcounta=3500000;
T=5;
elas=-.1;
pricecap=1000;
perspective=1;
printf("*****price cap = %f elasticity = %f perspective = %i sigma = %i
\n",pricecap,elas,perspective,sigma);
FP= fopen("diss.dat","a");
fprintf(FP,"*****price cap = %f elasticity = %f perspective = %i sigma = %i
\n",pricecap,elas,perspective,sigma);
fclose(FP);
mymain();

sigma=1;
maxcounta=3500000;
T=5;
elas=-.1;
pricecap=50;
perspective=2;
perspective2=1;
printf("*****price cap = %f elasticity = %f perspective = %i sigma = %i
\n",pricecap,elas,perspective,sigma);
FP= fopen("diss.dat","a");
fprintf(FP,"*****price cap = %f elasticity = %f perspective = %i sigma = %i
\n",pricecap,elas,perspective,sigma);
fclose(FP);
mymain();
} //main

```

```
//*****MAIN*****  
*****
```